Sistemas e Redes de Telecomunicações

João J. O. Pires

Departamento de Engenharia Electrotécnica e de Computadores
Instituto Superior Técnico

2006

O autor agradece que possíveis comentários sobre gralhas ou incorrecções no texto sejam enviados para: jpires@lx.it.pt

Índice

1 Introdução

1.1 Evolução das telecomunicações

A rede de telecomunicações que nos dias de hoje cobre o globo terrestre é sem dúvida a mais complexa, extensiva e cara de todas as criações tecnológicas, e porventura a mais útil de todas, na medida que constitui o sistema nervoso essencial para o desenvolvimento social e económico da civilização.

As telecomunicações são uma ciência exacta cujo desenvolvimento dependeu fortemente das descobertas científicas e dos avanços na matemática que tiverem lugar na Europa durante o século XIX. Foram as descobertas na área do electromagnetismo, que criaram as condições para o aparecimento do primeiro sistema de telecomunicações baseado na electricidade: o telégrafo.

O telégrafo foi patenteado no Reino Unido por Cooke e Wheatstone, em 1837. No entanto, o sistema por eles desenvolvido requeria cinco condutores metálicos, não sendo por isso muito prática a sua implementação.

Foi a criação por Morse do código que tem o seu nome, que veio dar o grande impulso à expansão do telégrafo. O primeiro sistema experimental orientado por Morse teve lugar nos Estados Unidos em 1844. Este sistema era claramente um sistema de transmissão digital, na medida em que a informação era transmitida usando pulsos de corrente. Tinha-se dois tipos de pulsos, um estreito (ponto) e outro mais longo (traço) e as diferentes letras eram codificadas através de combinações desses pulsos.

O sucesso do telégrafo foi tal que logo em 1866 foi instalado um cabo submarino transatlântico ligando o Reino Unido aos Estados Unidos. Em 1875, Portugal e o Brasil também ficaram ligados através de um outro cabo. Em 1875, a rede de cabos de telégrafo já cobria todo o globo incluindo o Extremo Oriente e a Austrália.

Outro ponto singular na grande caminhada das telecomunicações foi a invenção do telefone. Embora a história da invenção do telefone seja um tanto nebulosa, com vários inventores a requererem para si os louros dessa invenção, a primeira demonstração com sucesso de transmissão electrónica de voz inteligível foi realizada por Alexander Graham Bell em 1876, consagrada no histórico apelo de Bell para o seu assistente: "*Mr. Watson, come here, I want to see you*". Interessantes são também as palavras proféticas por ele proferidas nessa altura: "*This is a great day with me and I feel I have at last struck the solution of a great problem-and the day is coming when telephone wires will be laid on the houses, just like water or gas, and friends converse without leaving home".*

Destaque-se que na história do telefone surge também o nome do inventor português Cristiano Bramão que no ano de 1879 estabeleceu uma ligação telefónica entre Lisboa e Setúbal usando um aparelho por ele concebido [Sa99].

Inventado o telefone tratava-se de resolver o problema da ligação entre os interlocutores envolvidos numa ligação telefónica. A primeira solução consistiu na utilização de centrais telefónicas manuais. No entanto, com o aumento do número de linhas a utilização deste tipo de centrais tornou-se impraticável. Para além disso, tinha-se

o problema da falta de privacidade: as operadoras podiam ouvir facilmente as conversas entre os interlocutores.

Foi exactamente a falta de privacidade das centrais manuais que levou Strowger a inventar a primeira central telefónica de comutação automática, com patente concedida em 1891. Essa invenção compreendia dois elementos básicos:

1) Um dispositivo (disco) para ser usado pelo assinante que gerava sequências de pulsos de corrente correspondentes aos dígitos de 0 a 9;
2) Um comutador localizado na central telefónica, no qual um braço rotativo se movia passo-a-passo num arco semi-circular com dez contactos, cada um ligado a uma linha de assinante, sendo o movimento controlado pelos pulsos de corrente enviados pelo assinante.

As centrais automáticas Strowger tiveram uma grande expansão em todo o mundo e aplicação generalizada até aos anos 70. Embora a primeira central de comutação automática tivesse sido inaugurada em Lisboa (Central Telefónica da Trindade) em 1925, e a digitalização da comutação se tivesse iniciado em 1987, em 1994, ainda existiam, na rede telefónica nacional cerca de 160 000 linhas servidas por centrais Strowger. No campo da comutação automática será também de referir o papel dos CTT, que desenvolveu tecnologia própria, a qual foi introduzida na sua rede em 1956 (estações tipo ATU 52 e ATU 54) e se continuou a usar até ao início da digitalização na década de oitenta.

Outro marco importante na história das telecomunicações foi a demonstração por Marconi em 1894 da telegrafia/telefonia sem fios. Até 1910, as ondas rádio foram usadas essencialmente para transmitir sinais telegráficos. Porém, com a invenção em 1907 por De Forest da válvula termo-iónica, tornou-se possível a geração e modulação de portadoras eléctricas e a radiotelefonia começou a dar os primeiros passos. Progressos tecnológicos nesta área permitiram estabelecer em 1914 um serviço transatlântico de telegrafia sem fios, e realizar em 1926 a primeira ligação telefónica (1 canal de voz) entre os Estados Unidos e a Inglaterra. Inaugurava-se, assim, a competição entre os serviços de telecomunicações sem fios e os serviços baseados numa transmissão guiada, que tem sido uma constante até aos dias de hoje.

Muitas outras inovações vieram contribuir para que as telecomunicações se tornassem, neste início de século, uma das mais poderosas indústrias. No entanto, existe uma outra que não pode deixar de ser referida: a invenção do PCM (*Pulse Code Modulation*). Esta foi feita por Alec Reeves em França em 1936. Como, porém, a transmissão de um sinal de voz digitalizado requeria uma largura de banda mínima de 32 kHz, muito superior aos 3 kHz requeridos pelo correspondente sinal analógico, a implementação dos primeiros sistemas experimentais teve de esperar até que nos anos sessenta a tecnologia do estado sólido a permitisse concretizar.

Outros dois factos importantes que contribuíram para modelar as telecomunicações dos dias de hoje foram a proposta para usar as fibras de vidro (fibras ópticas) como meio de transmissão da informação, feita em 1966 por K. Kao e G. Hockman [Ka66] e o conceito de comutação de pacotes apresentado em meados da década de sessenta por P. Baran [Ba66]. A evolução da tecnologia de transmissão óptica permitiu que por exemplo nas últimas duas décadas a capacidade dos sistemas de transmissão aumentasse mais de 10 000 vezes, atingindo hoje em dias débitos superiores a 1 Tbit/s, e a comutação de

pacotes foi responsável pela enorme expansão das redes de dados e pelo consequente êxito da Internet.

Não se poderia deixar também de referir a expansão das comunicações móveis que ocorreu particularmente na última década e que foi fortemente impulsionada pela normalização do GSM[1] e também pela contribuição da engenharia de telecomunicações portuguesa com a concepção e implementação do pré-pago. Toda a infra-estrutura de rede inteligente que suporta o plano de controlo responsável por esse conceito, que hoje é usado por mais de mais de 80% dos telemóveis de todo o mundo, foi desenvolvido nos laboratórios da PT Inovação.

Depois deste breve percurso por alguns dos factos mais marcantes da evolução das telecomunicações, será pertinente colocar a questão: **o que são as telecomunicações?** Para começar, o prefixo tele é derivado do grego e significa "à distância". Poder-se-á, assim, dizer que as **telecomunicações compreendem o conjunto dos meios técnicos necessários para transportar e encaminhar tão fielmente quanto possível a informação à distância.** Esta definição deverá, no entanto, ser complementada com os seguintes comentários:

- Os principais meios técnicos são de natureza electromagnética;
- A informação a transmitir pode tomar diversas formas, nomeadamente, voz, música, imagens fixas, vídeo, texto, dados, etc.;
- Os sistemas de telecomunicações devem garantir um elevado grau de fidelidade, garantindo que a informação é transmitida sem perdas nem alterações;
- A fiabilidade é outra exigência primordial, já que o utilizador espera das telecomunicações um serviço permanente e sem falhas;
- O transporte da informação à distância é um problema da **transmissão**, que é um ramo importante das telecomunicações;
- Outro ramo importante é a **comutação**, que tem como objectivo o encaminhamento da informação;
- As redes de telecomunicações de hoje são máquinas de grande complexidade e, por isso, a sua **gestão e controlo** são tarefas de grande importância e claramente individualizadas das funções anteriores. A gestão de rede é responsável, nomeadamente, por garantir elevados graus de fiabilidade (resposta rápida a falhas) e flexibilidade (configuração rápida dos elementos de rede). O controlo da rede é realizado através da sinalização e é responsável pela dinâmica da rede e pela sua capacidade de resposta aos pedidos dos utilizadores.

[1] *Global System for Mobile Communications*

1.2 Normalização em Telecomunicações

Devido ao carácter internacional das telecomunicações é fundamental a normalização, sobretudo, em certos aspectos mais relevantes tais como:

- **Aspectos técnicos**: definição da qualidade de serviço e dos parâmetros que a influenciam; especificação das interfaces, nomeadamente, dos sinais usados na transmissão e sinalização, etc.
- **Planificação geral da rede**: estrutura da rede internacional, plano de transmissão, distribuição dos números telefónicos, etc.
- **Problemas de exploração e gestão**: definição dos preços das chamadas internacionais, análise do tráfego, etc.

No plano das redes nacionais a normalização também é importante de modo a:

- Garantir a compatibilidade dos sistemas provenientes de fabricantes diferentes;
- Assegurar a mesma qualidade de serviço mínima a todos os utilizadores;
- Respeitar as convenções internacionais.

O principal organismo de normalização na área das telecomunicações é a ***International Telecommunication Union*** (ITU). Este organismo é uma agência da ONU e actua fundamentalmente através dos seguintes orgãos:

- **ITU *Telecommunication Sector*** (ITU-T), que corresponde ao antigo "*Comité Consultatif International Télégraphique et Téléphonique*" (CCITT). As suas funções incluem o estudo de questões técnicas, métodos de operação e tarifas para as comunicações telefónicas e de dados.
- **ITU *Radiocommunications Sector*** (ITU-R), que corresponde ao antigo "*Comité Consultatif International des Radiocommunications*" (CCIR). Estuda todas as questões técnicas e operacionais relacionadas com radio-comunicações, incluindo ligações ponto-a-ponto, serviços móveis e radiodifusão. Associado ao ITU-R está o ***International Frequency Registration Board*** (IFRB), que regula a atribuição das bandas de frequências aos diferentes serviços.

Tanto a ITU-T como a ITU-R são compostos de delegados dos governos, operadores de telecomunicações e organizações industriais. Ambos têm um elevado número de grupos de estudo. Os seus trabalhos tomam a forma de recomendações, que são ratificadas por assembleias plenárias, que têm lugar de quatro em quatro anos. Os resultados dessas sessões plenárias são publicados numa série de volumes, que proporcionam recomendações e informação actualizada para todos os interessados na área das telecomunicações.

Além da ITU, existem um conjunto de outros organismos também com actividade na área das telecomunicações. Refira-se, por exemplo, a *International Standards Organization* (ISO). Este organismo tem índole mundial e actividade de normalização em diferentes áreas, incluindo as tecnologias de informação. Recorde-se que, o modelo OSI (*Open Systems Interconnect*), o conhecido protocolo de sete camadas que regula as

comunicações entre computadores, foi definido por esta organização. O ANSI (*American National Standards Institute*) que é o corpo americano da ISO também tem produzido inúmeras contribuições de relevo. Foi esta agência a responsável pela norma ASCII (*American Standard Code for Information Interchange*), usada por muitos computadores para armazenar informação.

Outro organismo digno de realce é o ETSI (*European Telecommunication Standards Institute*), o qual foi criado em 1988 para desenvolver as normas necessárias à criação de uma rede de telecomunicações pan-europeia. Esta instituição tem tido um papel importante no desenvolvimento das normas para as redes móveis e em particular da norma GSM (G*lobal Systems for Mobile Communications*). O IEEE (*Institute of Electrical and Electronic Engineers*) é uma associação profissional do tipo da Ordem dos Engenheiros, que também tem contribuído com inúmeras normas para o progresso das telecomunicações. Por exemplo, as normas que delinearem a Ethernet (IEEE 802.3) e as normas que têm projectado o seu desenvolvimento foram concebidas por esta associação.

Actualmente, também existem também inúmeros consórcios de fabricantes e operadores dirigidos a determinadas áreas específicas, com o objectivo de desenvolverem normas de equipamentos ou serviços em curtos espaços de tempo. Veja-se por exemplo o Fórum DSL (www. dslforum.org) ou o Fórum ATM (www. ATMforum.org).

1.3 Conceitos fundamentais e topologias

As redes de telecomunicações são constituídas por uma infinidade de equipamentos baseados numa grande diversidade de tecnologias e em muitos casos concebidos e instalados em épocas muito diferentes. Por exemplo, na rede telefónica pública comutada portuguesa há cerca de dez anos atrás ainda era possível encontrar em paralelo com as modernas centrais de comutação digitais, centrais de comutação analógicas com várias dezenas de anos.

Numa rede de telecomunicações podem-se identificar dois tipos básicos de equipamentos: **vias de transmissão** ou canais de transmissão e elementos (ou dispositivos) de rede designados genericamente por **nós.** As vias de transmissão asseguram a transmissão da informação e a interligação entre os diferentes nós. As vias de transmissão podem ser simples pares de condutores de cobre (pares simétricos) como é o caso da linha telefónica até meios de transmissão mais complexos como é o caso das fibras ópticas. Não se pode igualmente ignorar o cabo coaxial usado nas redes de distribuição de televisão e os canais via rádio usados nas redes celulares ou nas redes de comunicação via satélite.

Os elementos de rede englobam nomeadamente o equipamento de comutação, o equipamento terminal, os servidores e os sistemas de sinalização e de gestão. O equipamento de comutação inclui por exemplo as centrais de comutação nas redes telefónicas ou os *routers* nas redes de dados e tem por objectivo assegurar o encaminhamento apropriado da informação. Nas redes telefónicas a comutação é feita usando **comutação de circuitos**, ou seja, antes do início da conversação estabelece-se por intermédio de sinalização uma ligação bidireccional entre os utilizadores intervenientes (circuito). Nas redes de dados a informação é fragmentada em pacotes, aos quais se adicionam um cabeçalho, que contem entre outra informação o endereço do destinatário.

O encaminhamento é feito com base na informação contida no cabeçalho e por conseguinte a técnica designa-se por **comutação de pacotes.**

Os equipamentos terminais servem de interface entre a rede e o utilizador e têm por objectivo processar a informação veiculada pela rede de modo apresentá-la numa forma inteligível ao utilizador. Estes equipamentos podem apresentar diferentes graus de complexidade, podendo ir deste um simples telefone, a um PPCA (Posto Particular de Comutação Automática)[2], passando por um computador pessoal. Os servidores são dispositivos que armazenam informação que pode ser difundida de modo livre ou a pedido dos utilizadores. Como exemplo deste tipo de elementos de rede têm-se os servidores WWW usados na Internet e as cabeças de redes das redes de televisão por cabo onde são transmitidos os diferentes canais que chegam a casa do utilizador.

Os sistemas de sinalização e gestão são responsáveis por processarem a informação de sinalização e gestão da rede. A informação de sinalização é responsável pela componente dinâmica das redes de telecomunicações e a informação de gestão pode corresponder a várias funções tais como detecção de falhas, configuração da rede, autorização de acesso, etc.

Nos últimos trinta anos surgiram um grande número de novos tipos de redes de telecomunicações e certamente que muito mais irão surgir no futuro. Embora seja importante para um engenheiro de telecomunicações ter um conhecimento detalhado das tecnologias usadas para construir essas diferentes redes, e compreender as suas potencialidades e limitações, também não deixa de ser relevante ser capaz de perspectivar essas redes em termos de entidades abstractas, independentes das tecnologias e assentes em conceitos gerais. Nesse sentido uma rede de telecomunicações pode-se representar através de um grafo. Um grafo é definido geometricamente como um conjunto de pontos designados por vértices interligados por um conjunto de linhas. Um grafo pode-se representar formalmente por $G=(V,E)$, onde $V=\{v_1,v_2,.....v_n.\}$ representa o conjunto dos vértices e $E=\{e_1,e_2,.....e_m.\}$o conjunto de linhas. A Figura 1.1 mostra, por exemplo, o grafo $G=(\{v_1,v_2......v_6\}\{e_1,e_2,.....e_8\})$. Para representar uma rede de telecomunicações através de um grafo faz-se corresponder um vértice a cada nó e uma linha a cada via de transmissão, como se mostra na Figura 1.2.

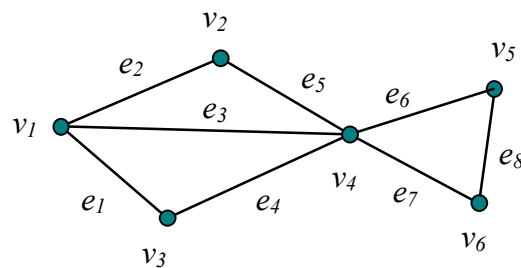



Figura 1.1 Exemplo de um grafo.

A estratégia de interligação entre os nós designa-se por **topologia da rede,** ou de um modo mais preciso por topologia física. Este refinamento na definição ajuda a distinguir o

[2] Segundo a designação anglo-saxónica PABX (*Private Automatic Branch Exchange*).

aspecto físico do modo como a informação é distribuída na rede que define a topologia lógica. Tendo como exemplo a Figura 1.2 pode-se admitir que o nó 1 funciona como nó distribuidor e que toda a comunicação é feita directamente entre os diferentes nós e o nó 1. Como consequência a **topologia física** e a **topologia lógica** são diferentes como se evidência na Figura 1.3 através da representação dos grafos correspondentes.

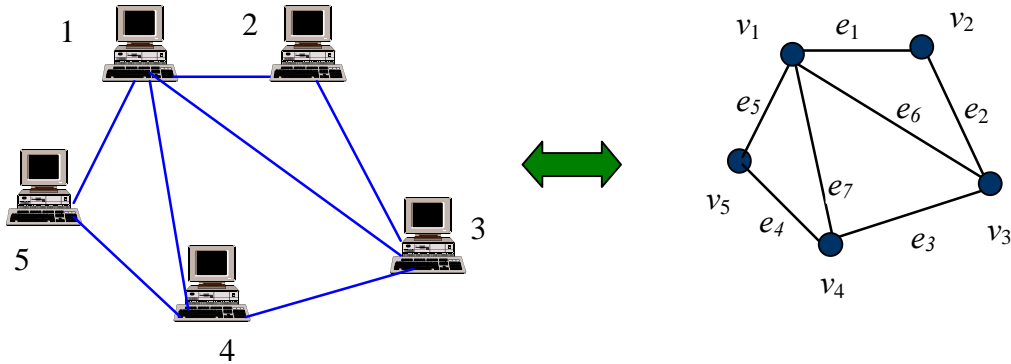



Figura 1.2 Exemplo de uma rede e do seu grafo equivalente.

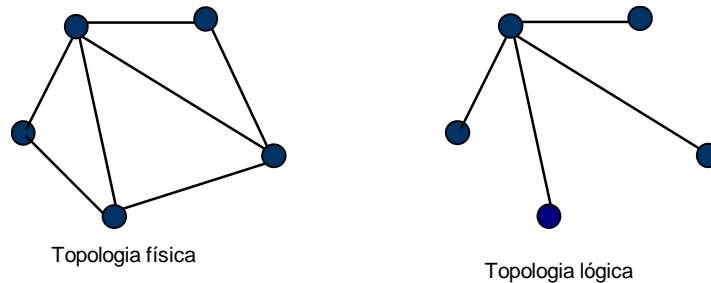



Figura 1.3 Comparação entre a topologia física e lógica.

Nas de redes de telecomunicações encontra-se uma grande variedade de topologias. Essas topologias condicionam à partida a estratégia de desenvolvimento e o tipo de serviços que a rede pode fornecer e por isso a definição adequada da topologia constitui uma etapa importante no processo de planeamento da rede. A topologia mais simples é a topologia em barramento (*Bus*). Como se mostra na Figura 1.4 nessa topologia a via de transmissão é partilhada por todos os elementos de rede. O facto do meio ser partilhado dificulta significativamente o processo de comunicação e impõe a existência de um protocolo de comunicação de modo a evitar colisões entre os sinais enviadas simultaneamente pelos diferentes nós. A **topologia em barramento** (*Bus*) é muito usada nas redes Ethernet nas quais aparece normalmente associada ao protocolo CSMA/CD (*Carrier Sense Multiple Access with Collision Detection*) (para detalhes sobre este protocolo ver referência [Sh99]).

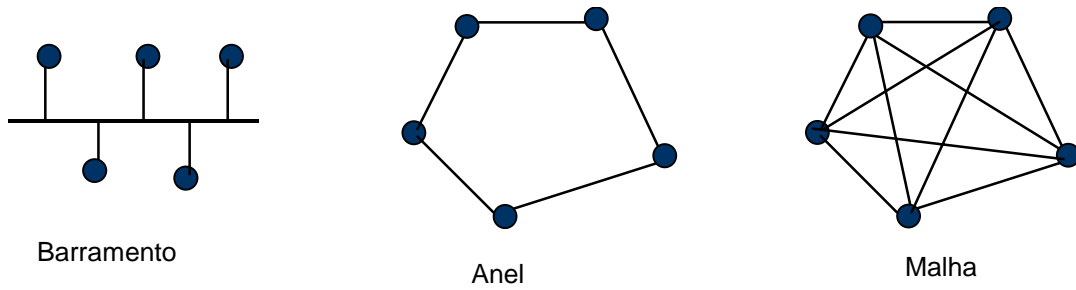



Figura 1.4 Grafos correspondentes a diferentes tipos de topologias.

Na **topologia em anel** cada nó só está interligado aos nós vizinhos. No caso de querer comunicar com outros nós as mensagens terão de ser enviadas através dos vizinhos. Uma rede em anel pode ser **unidireccional** ou **bidireccional**. No caso unidireccional toda a informação viaja no mesmo sentido e cada nó só pode comunicar directamente com um vizinho, enquanto no caso bidireccional a informação viaja nos dois sentidos e cada nó pode comunicar directamente com os dois vizinhos. Esta topologia começou por ser usada nas redes de computadores (veja-se o caso das redes *token ring e* FDDI[3]), mas hoje é popular em muitos outros ambientes, como sejam por exemplo as redes baseadas na hierarquia digital síncrona ou as redes RPR (*Resilient Packet Ring*) (para detalhes sobre estas redes ver a referência [Gu03]). Essa popularidade advém do facto desta topologia permitir garantir graus de fiabilidade elevados (resistência a falhas) com um consumo modesto de recursos de transmissão.

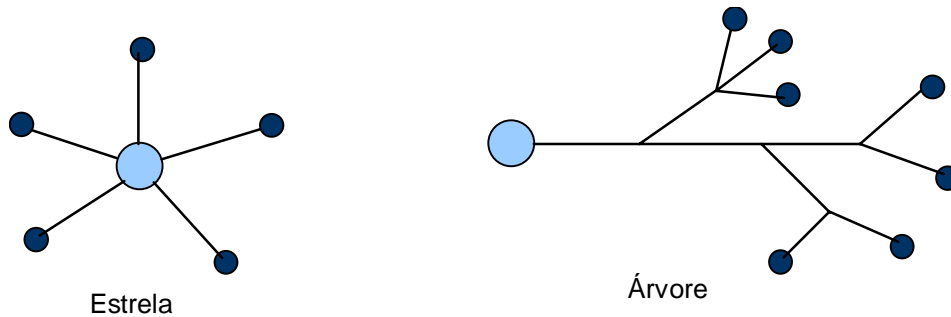



Figura 1.5 Grafos correspondentes às topologias em anel e árvore.

A **topologia em malha** é uma topologia com conexão total caracterizada por apresentar uma ligação directa entre todos os pares de nós. Numa rede baseada nesta topologia o processo de comunicação está muito facilitado, pois qualquer troca de informação entre dois nós não envolve a intervenção de mais nenhum outro nó. A principal desvantagem desta solução reside na grande quantidade de recursos de transmissão que exige. Por exemplo, uma rede com N nós baseada nesta topologia requer $N(N-1)/2$ vias de transmissão. Quando N>>1, esse número é aproximadamente proporcional a N^2, o que faz com que essa topologia se torne impraticável quando o valor

[3] *Fiber Distributed Data Interface*

de N ultrapassa algumas dezenas de nós. Essa inexequibilidade é evidente se nos situarmos, por exemplo, na componente de acesso de uma rede telefónica, onde um número típico de 10^4 utilizadores exigia a utilização de 50 milhões de linhas telefónicas. Apesar destas limitações a malha é muito usada para interligar os nós principais das redes de telecomunicações, já que permite garantir um grau de disponibilidade muito elevado[4].

A **topologia em estrela** é a solução, normalmente usada sempre que é necessário interligar um elevado número de nós. Nesta topologia há uma diferenciação entre as funcionalidades do nó central e as dos restantes nós, já que é este nó que controla as comunicações entre todos os outros. A existência de um nó com responsabilidade acrescida, indica que nesta topologia o controlo do processo de informação é centralizado. A solução em estrela tem sido muito usada nas redes telefónicas, particularmente na rede de acesso, onde todo o fluxo de informação com o utilizador é controlado pelo comutador local, permitindo concentrar o equipamento sofisticado e caro na rede, e garantir que o equipamento terminal usado pelo utilizador é relativamente simples.

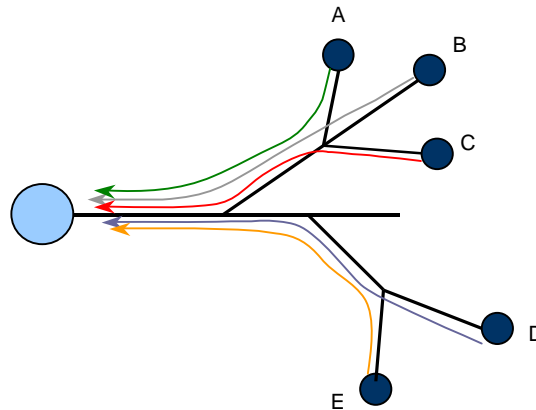



Figura 1.6 Colisão entre os sinais gerados por cinco utilizadores numa rede em árvore.

A **topologia em árvore** surgiu associada a serviços distributivos, onde o objectivo é difundir o mesmo sinal desde o nó onde é gerada para todos os outros nós. Esta perspectiva distributiva faz com que nos vários pontos de divisão o sinal seja repetido até atingir o equipamento terminal do utilizador. Esta solução foi desenvolvida no âmbito das redes de distribuição de televisão por cabo, também referidas na literatura como redes de cabo. Com o desenvolvimento tecnológico foi possível introduzir nestas redes um canal ascendente, garantindo, assim, a bidireccionalidade requerida pelos serviços interactivos tais como serviço telefónico e o acesso à Internet. Será, no entanto de destacar, que nestas redes o meio de transmissão também é partilhado por todos os utilizadores, perdendo-se, assim, a simplicidade característica do anel. Como consequência, surge tal como nas redes em barramento o problema da colisão entre os sinais enviados pelos vários utilizadores (ver Figura 1.6) e a necessidade de usar algoritmos de contenção apropriados para regular o processo de comunicação.

[4] Uma rede em anel permite responder a um único corte numa ligação, enquanto uma rede em malha permite responder a vários cortes.

1.4 Arquitectura de rede

1.4.1 Plano de rede

Em geral, numa rede de telecomunicações podem-se individualizar três entidades como funcionalidade distintas. Essas entidades designam-se por planos de rede e consistem no **plano de utilizador, plano de controlo** e **plano de gestão**, como se mostra na Figura 1.7.

O **plano do utilizador** é responsável por transferir a informação do utilizador através da rede. Essa transferência de informação faz uso, nomeadamente, do equipamento terminal, dos dispositivos de rede e das vias de transmissão. De modo algo simplista pode-se dizer que este plano é responsável por assegurar o suporte físico às ligações solicitadas pelos utilizadores.

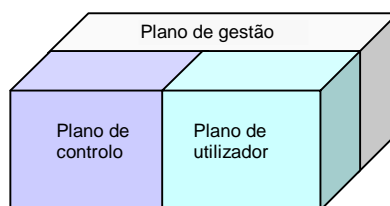



Figura 1.7 Planos de rede.

O **plano de controlo** pode ser visto como o sistema nervoso da rede, adicionando ao plano do utilizador a dinâmica necessário para poder responder aos requisitos do utilizador. Este plano é por exemplo responsável pelo processo de sinalização associado ao estabelecimento, supervisão e terminação das ligações. É este plano que é responsável pela actualização das tabelas de encaminhamento dos *routers,* ou pela geração dos sinais de controlo necessários actuar sobre os comutadores telefónicos. O sinal de toque, de linha ou de linha ocupada são exemplos nossos conhecidos de sinais gerados pelo plano de controlo. Um exemplo típico de um plano de controlo é o Sistema de Sinalização nº 7 usado nas redes de telefonia fixa e celulares.

O **plano de gestão** tem vindo a ter a ter uma importância cada vez maior nas redes de telecomunicações, pois é fundamental para garantir elevados graus de fiabilidade (tempos de detecção e correcção de falhas muito reduzidos) e flexibilidade (reconfiguração rápida dos elementos de rede de modo a responder por exemplo a picos de tráfego). As funções asseguradas por este plano são muito diversas, podendo referir-se, entre outras, funções ao nível de detecção, diagnóstico e correcção de falhas (gestão de falhas), funções ao nível da modificação das configurações da rede (gestão de configuração), funções ao nível de monitorização do desempenho da rede (gestão de desempenho) e funções ao nível de identificação, validação e autorização de acesso à rede (gestão de segurança). Exemplificando: um corte de um cabo vai gerar uma série de alarmes que vão ser analisados pelo sistema de gestão, através da funcionalidade de gestão de falhas, de modo a dar uma resposta adequada. Essa resposta pode ser, por exemplo, a geração de uma alarme que avise o operador do sistema da ocorrência do problema e ainda a reconfiguração da rede de modo a gerar percursos alternativos para o tráfego afectado pela falha. Note-se que, as funções deste plano aparecem muitas vezes associada à abreviatura OAM, que corresponde às iniciais de operação, administração e manutenção.

Grosso modo as componentes de operação e manutenção correspondem às funções de gestão de falhas, de configuração e de desempenho, enquanto a componente de administração corresponde à função de gestão de segurança.

1.4.2 Estratificação em camadas

As redes de telecomunicações são estruturas altamente complexas, consistindo em várias estruturas heterogéneas realizando uma grande diversidade de funções. Para simplificar a sua concepção, desenvolvimento e operação é usual recorrer-se a uma variação do velho princípio de "dividir para conquistar". Por outras palavras, o conjunto das funções da rede é dividido em grupos designados por **camadas** (*layers*). Cada camada assume a responsabilidade por um certo número de tarefas próprias, de modo que quando todas as camadas operam em conjunto, é criada uma rede completa e funcional. Outra vantagem da estratificação em camadas é a de permitir uma evolução gradual e suave da rede, possibilitando que cada camada evolua de modo independente e possa, assim, capitalizar a introdução de novas tecnologias específicas para essa camada.

Para que a estratificação em camadas possa ser benéfica é necessário especificar sem ambiguidades os serviços proporcionados pelas diferentes camadas e as interfaces entre elas. Como exemplo dessas especificações refira-se a recomendação X.200 do ITU-T, onde a arquitectura do modelo OSI (*Open Systems Interconnection*) é definida. Como os leitores com conhecimentos na área das redes de computadores sabem, esse modelo define uma arquitectura com sete camadas que regula as comunicações entre computadores.

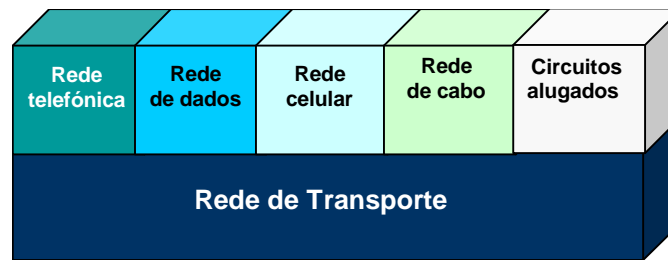



Figura 1.8 Estrutura estratificada de uma rede de telecomunicações.

Nas redes que fazem uso de diferentes tecnologias de rede[5] também é útil usar uma representação em camadas de modo a ajudar a visualizar a rede de um modo simples e facilmente compreensível. Neste caso, a cada tecnologia de rede faz-se corresponder uma camada de rede, sendo que as camadas sucessivas estabelecem entre si uma relação tipo cliente-servidor. Tendo em conta esta perspectiva é usual dividir-se uma rede de telecomunicações em **camada de rede de serviço** e **camada de rede de transporte**. A camada de rede de serviço consiste em diferentes redes de serviço, cada uma responsável por um certo tipo de serviço. Como exemplo, refira-se o serviço de comutação de circuitos telefónicos, o serviço de comutação de pacotes, o serviço de linhas alugadas, etc. (ver Figura 1.8).

[5] Como exemplo de tecnologias de rede refira-se o IP (*Internet Protocol*), o ATM (*Asynchronous Transfer Mode*), SDH (*Synchronous Digital Hierarchy)*, WDM (*Wavelength Division Multiplexing)*.

A camada de rede de serviço funciona como cliente da camada de rede de transporte e é responsável por recolher, agregar e introduzir a informação (voz, dados, multimédia, etc.) nessa rede. A camada de rede de transporte pode ser vista como uma plataforma tecnológica que assegura uma transferência transparente, fiável e independente dos serviços, da informação à distância, garantindo funcionalidades tais como transmissão, multiplexagem, encaminhamento, protecção e supervisão e aprovisionamento de capacidade. Essa rede é constituída por diferentes elementos de rede interligados segundo uma certa topologia física (normalmente anel ou malha) e interagindo directamente com o plano de gestão. O se objectivo é proporcionar conexões, designadas aqui por **caminhos**, à camada de rede de serviço. Para a camada de rede de serviço os caminhos são vistos como ligações físicas entre os seus elementos da rede. Tomando por analogia a rede rodoviária, a camada de rede de transporte corresponde à rede de auto-estradas, enquanto a camada de rede de serviço corresponde aos diferentes tipos de veículos (carros, autocarros, camiões, etc.) que circulam nas auto-estradas.

A principal função da camada de rede de transporte consiste em proporcionar caminhos aos seus clientes (rede telefónica comutada, rede RDIS, rede IP, rede ATM, etc). O **caminho** é uma ligação bidireccional semi-permanente extremo-a-extremo estabelecida por acção do plano de gestão de rede, ou manualmente. Convencionalmente, distingue-se do **circuito** na medida em que este corresponde a uma ligação bidirecional extremo-a-extremo, estabelecida e terminada dinamicamente por acção da sinalização enviada pelo utilizador, na base de uma provisão de curta duração, como acontece, por exemplo, com as ligações telefónicas comutadas. Contudo, hoje em dia, com os novos conceitos da rede ASTN (*Automatic Switched Transport Network*) essas distinções tornam-se algo mais difusas.

> Note que em muitos casos na literatura da área não há uma distinção clara entre os conceitos de caminho e circuito. No âmbito desta disciplina considera-se que são conceitos diferentes: o caminho é estabelecido por acção do plano de gestão, enquanto o circuito é estabelecido pela acção do plano de controlo.

A Figura 1.9 exemplifica o conceito da estratificação de uma rede de telecomunicações. Nessa figura a rede de transporte, representada pelo plano inferior, é constituída por cinco multiplexadores de inserção/extracção ou ADMs (*add/drop multiplexer*) (ADMs A, B, C, D e E) os quais são interligados por fibras ópticas formando uma topologia física em anel. A funcionalidade desses elementos de rede irá ser estudada posteriormente no capítulo 5. Por sua vez, a rede de serviço considerada é uma rede telefónica comutada cujos elementos de rede são centrais de comutação (CC). Como se pode ver, a ligação entre as centrais d e c é estabelecida através do caminho criado pelos ADMs E, D e C. De modo similar, os caminhos B-C e A-B-C vão interligar as centrais b-c e a-c, implementando uma topologia lógica em estrela, com a central c a funcionar como nó agregador. Do ponto de vista da rede telefónica somente a topologia lógica da camada de serviço é visível, sendo independente do modo como os caminhos estão estabelecidos.

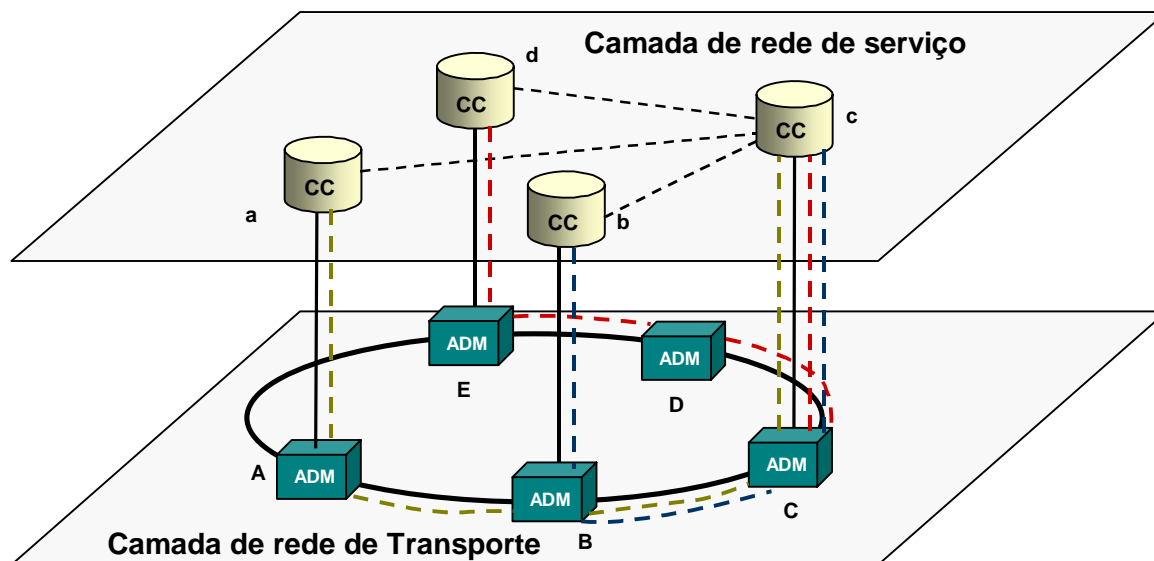



Figura 1.9 Exemplificação do papel da camada de transporte.

As tecnologias usadas na implementação rede de transporte são completamente independentes do serviço. Essas tecnologias são baseadas fundamentalmente na SDH e na WDM, as quais irão ser abordadas no Capítulo 5. A SDH é usada para estabelecer caminhos no domínio eléctrico[6], enquanto a WDM é usada para a mesma função no domínio óptico. As redes de serviço tradicionalmente estão ligadas a tipos de serviços bem definidos, embora esta identificação simplista em muitos casos já não faça hoje sentido. Na Secção 1.5 irão ser analisadas algumas das principais redes de serviço.

1.4.3 Hierarquização da rede

As redes de telecomunicações são normalmente segmentadas numa estrutura hierárquica com três níveis: **núcleo** ou backbone, **metro**, e **acesso**. A rede de núcleo abarca grandes distâncias (centenas a milhares de quilómetros) é usada para transportar grandes agregados de tráfego e oferecer conectividade às redes regionais ou metropolitanas. No oposto da hierarquia situam-se as redes de acesso que são concebidas para oferecer directamente conectividade a uma grande variedade de utilizadores. As redes de acesso usam uma grande variedade de tecnologias e protocolos de comunicação que dependem dos serviços proporcionados ao utilizador. Por exemplo, se o utilizador requer apenas o serviço telefónico a rede de pares simétricos convencional é suficiente, mas se o utilizador desejar um acesso de banda larga para ligações rápidas à Internet ou para serviços de vídeo é já necessário equacionar-se outras alternativas como o ADSL, redes de cabo coaxial, ou mesmo acesso em fibra óptica. Situadas no meio estão as redes metropolitanas ou regionais. Estas estendem-se em média por regiões entre 10-100 km e interligam o acesso e o núcleo.

[6] Os débitos típicos dos caminhos no domínio eléctrico são 34 Mb/s e 155.5 Mb/s.

Como se mostra na Figura 1.10 a topologia típica de uma rede de núcleo é a malha, enquanto da rede metropolitana é normalmente baseada numa topologia em anel. Já no acesso encontra-se uma maior variedade de topologias, podendo dizer-se que com excepção da malha todas as topologias estudadas na Secção 1.3 estão aqui representadas.

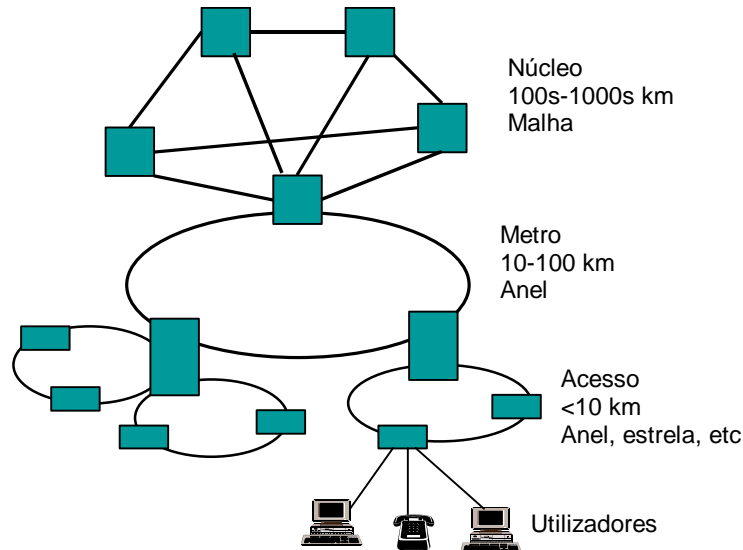



Figura 1.10 Hierarquia de uma rede de telecomunicações.

A estrutura hierárquica definida é transversal às redes de transporte e de serviços. Por exemplo, na rede de serviço usada para o transporte de dados (rede de dados) a componente do acesso tem a designação de LAN (*Local Area Network)*, a componente metro designa-se por MAN (*Metropolitan Area Network*) e a componente do núcleo por WAN (*Wide Area Network*).

Note-se que na rede de núcleo e metropolitana a topologia física da rede é imposta pela camada de transporte, a qual vai proporcionar a conectividade necessária para garantir às camadas de serviço as topologias lógicas requeridas. No que diz respeito ao acesso as camadas de serviço têm um papel importante na definição da topologia física. Por exemplo no acesso telefónica a topologia típica é a estrela, no cabo é a árvore e nas LANs é o barramento ou estrela.

1.5 Tipos de redes de serviço

1.5.1 Rede telefónica pública comutada

Numa rede telefónica convencional o equipamento terminal é essencialmente o telefone. As vias de transmissão incluem o meio de transmissão (cabos de pares simétricos, fibra óptica, ondas hertzianas, etc.) e os repetidores. O equipamento de comutação é constituído pelos comutadores digitais localizados nas centrais de comutação telefónica, que irão ser estudados no Capítulo 6.

1.5.1.1 Estrutura geral

A estrutura mais simples que é possível conceber para uma rede telefónica comutada está representada na Figura 1.11 e consiste numa central de comutação telefónica directamente ligada ao equipamento terminal dos utilizadores através de uma linha telefónica (**central local**) de acordo com uma topologia física em estrela.

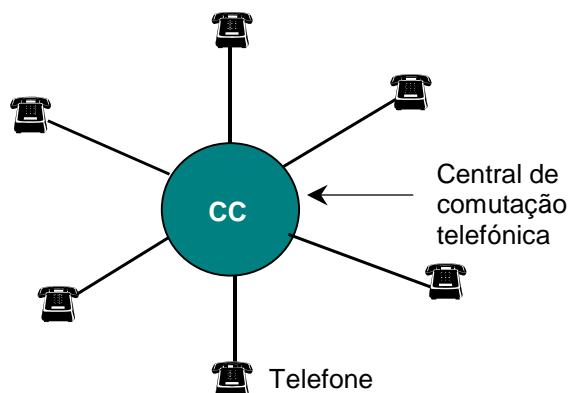



Figura 1.11 Rede em estrela com comutação centralizada.

Quando a área coberta pela rede em estrela e o número de utilizadores por ela servidos cresce, o preço da linha telefónica aumenta. Então, torna-se mais económico dividir essa rede em várias redes de pequenas dimensões, cada uma delas servida pela sua própria central de comutação telefónica. Nesse caso, o comprimento médio da linha de assinante decresce, diminuindo o seu custo total, mas em contrapartida o custo associado à comutação aumenta. Como se exemplifica na Figura 1.12 existe um número de centrais telefónicas óptimo, para o qual o custo total é mínimo.

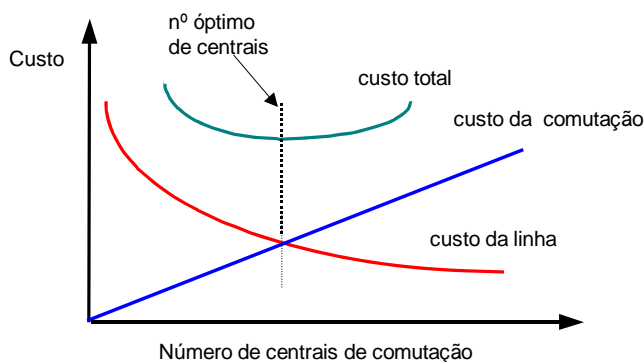



Figura 1.12 Variação do custo da rede com o número de centrais.

Numa área servida por diferentes centrais locais, os utilizadores de uma central terão certamente necessidade de comunicar com os utilizadores de outras centrais. É, assim, necessário estabelecer ligações, ou junções, entre as diferentes centrais, formando-se a **rede de junção**. Se as junções são estabelecidas entre todas as centrais locais, tem-se uma

rede de junção com conexão total baseada numa topologia física em malha. Porém, uma rede em malha neste nível da rede pode não ser económica, sendo preferível em muitos casos ligar as centrais entre si através de um centro de comutação central, designado por **central tandem** (Figura 1.13). Note-se que, a presença desta central introduz uma estrutura hierárquica na rede.

Na prática, junções directas entre centrais são economicamente justificáveis, quando se tem um grande fluxo de tráfego, ou quando as distâncias são curtas. Inversamente, quando o tráfego é reduzido e as distâncias são grandes é preferível o encaminhamento indirecto através de uma central tandem.

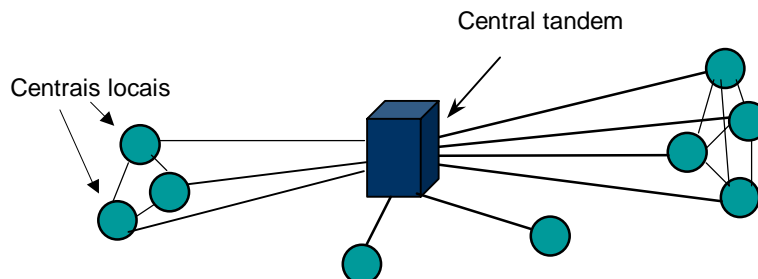



Figura 1.13 Área servida por várias centrais de comutação.

Os utilizadores da rede, para além das ligações locais, necessitam de comunicar com outros utilizadores localizados em diferentes áreas de um país. As diferentes áreas estão ligadas entre si por circuitos de longa distância, que constituem a **rede de núcleo,** também designada na gíria telefónica por **rede de troncas** ou de **longa distância**. Tal, como não é económico as centrais locais estarem todas ligadas entre si, também não é muitas vezes económico ter as centrais de longa distância totalmente interligadas. Assim, surgem os **centros de trânsito**, para encaminhar o tráfego entre as diferentes áreas, fazendo com que uma rede telefónica nacional apresente uma estrutura hierárquica, como se exemplifica na Figura 1.14. Em termos de topologia, essa rede apresenta uma topologia em árvore não pura, já que, quando se sobe na hierarquia aumenta, o número de ligações entre centrais do mesmo nível hierárquico. Um centro de trânsito primário constitui a interface entre a rede de junção e a rede de núcleo. Cada central local está ligada a um centro primário, seja directamente, seja através de uma central de junção tandem.

Um centro primário constitui a primeira camada da rede de núcleo, sendo o número de camadas adicionais dependente da dimensão do país. A Figura 1.14 mostra uma rede de núcleo constituída por duas camadas. Neste caso, a camada mais elevada é a segunda, sendo caracterizada por uma topologia em malha[7], com cada centro de comutação telefónica ligado directamente à central internacional do país.

Em síntese, uma rede telefónica nacional é baseada numa estrutura hierárquica constituída pela interligação das seguintes redes:

1) Rede privada de utilizador. Consiste numa rede dentro das instalações do utilizador e pode ser, por exemplo, constituída por várias linhas telefónicas, ligando equipamento terminal a uma central PPCA.

[7] Está-se a referir a uma topologia lógica, pois a conectividade física é determinada pelo transporte.

2) Rede de acesso ou local, que é responsável por ligar os telefones ou PPCAs dos assinantes à central local. Embora não representados na Figura 1.14 a rede de acesso pode também incluir concentradores, como se irá ver posteriormente.
3) Rede de junção, que interliga um grupo de centrais locais e ligando-as por sua vez ao centro de trânsito primário.
4) Rede de núcleo ou rede de troncas, que interliga os centros de trânsito[8] através do país.

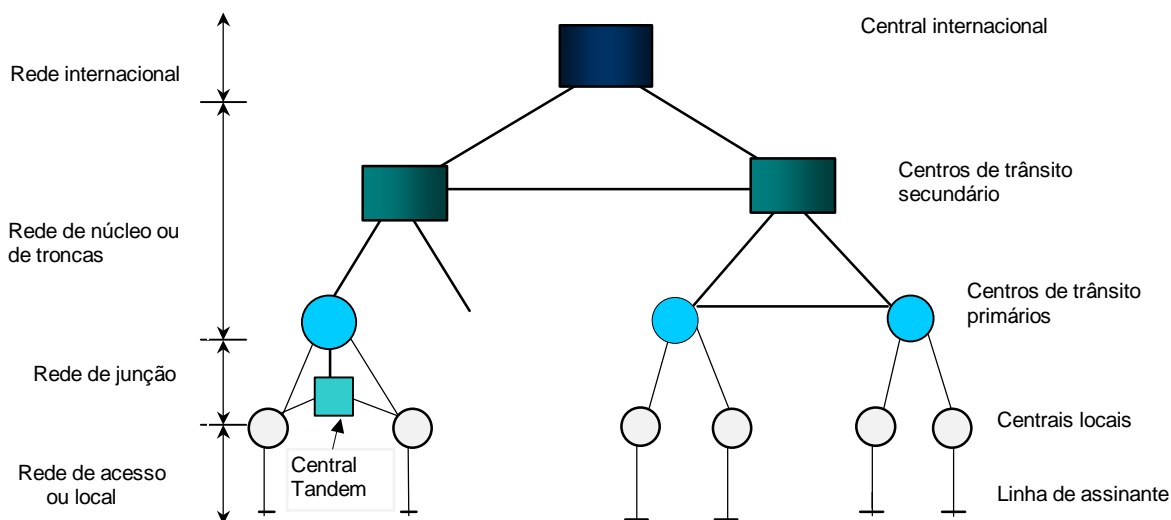



Figura 1.14 Organização hierárquica de uma rede telefónica nacional.

Note-se que, segundo a terminologia da ITU-T a rede local é constituída pelo conjunto da rede de acesso e rede de junção. Na generalidade dos casos, usa-se no âmbito desta disciplina a definição de rede local apresentada no ponto 2.

Embora a componente de acesso das redes telefónicas seja objecto de estudo detalhado no Capítulo 4, convém referir que uma parte muito significativa dessa componente é constituída pela infra-estrutura de cobre que liga o telefone dos utilizadores às centrais locais, fazendo com que a rede telefónica seja conhecida, particularmente nos meios jornalísticos, pela designação algo imprópria de rede de cobre. A infra-estrutura de cobre é implementada recorrendo a fios de cobre isolados e entrelaçados designados por pares simétricos. As enormes limitações dos pares simétricos, nomeadamente a sua largura de banda muito reduzida e a enorme susceptibilidade às interferências, têm levado os engenheiros de telecomunicações a explorar outras soluções, sobretudo quando está em causa o acesso de banda larga. A solução mais sólida é sem dúvida a que faz uso da fibra óptica, sendo de referir entre outras as seguintes alternativas: ligação em fibra óptica entre a central local e um armário exterior, ligando em seguida o armário às instalações do utilizador através da par simétrico (fibra até ao quarteirão); ligar a fibra directamente até às instalações do utilizador (fibra até casa). Ainda que a primeira solução seja uma

[8] O centro de trânsito primário também é designado por centro de grupo de redes e o centro de trânsito secundário por estação distribuidora.

alternativa mais apropriada para implementação a curto prazo, faz-se notar que há já um número significativo de operadores de telecomunicações em vários países a apostar na última solução que é sem dúvida a mais arrojada.

Uma outra singularidade das redes de acesso reside no facto de usar **transmissão a dois fios**, ou seja nas comunicações aí estabelecidas o sinal telefónico transmitido e o recebido partilharem a mesma linha telefónica. Na transmissão de dados (por exemplo em ADSL) tal situação coloca problemas graves, sendo necessário usar técnicas de duplexagem para separar os dois sentidos de transmissão. Outro problema prende-se com a amplificação e regeneração dos sinais. Os amplificadores e regeneradores bidireccionais não são soluções práticas. Assim, nos casos em que é necessário processar o sinal durante a transmissão é necessário separar fisicamente os dois sentidos de comunicação - **transmissão a quatro-fios**. É esse o caso da transmissão nas componentes de metro, de núcleo e internacionais das redes de telecomunicações. Será ainda de referir que os comutadores digitais das centrais de comutação também exigem uma transmissão a quatro fios. A conversão de uma transmissão de dois para quatro fios ou vice-versa é feita usando um dispositivo designado por **híbrido,** cuja estrutura e funcionalidades irão ser detalhadas no Capítulo 5.

1.5.1.2 Critérios de qualidade e plano de transmissão

Atendendo ao carácter internacional das telecomunicações, qualquer rede nacional deve obedecer a critérios de qualidade bem definidos, de modo a que, a qualidade dos circuitos estabelecidos nas ligações internacionais seja tão independente quanto possível do percurso entre o emissor e o receptor da informação.

No planeamento das redes de telecomunicações analógicas[9] um dos aspectos mais importantes era o controlo da atenuação. Dentro desse enquadramento foi definido um parâmetro – o **equivalente de referência** - que foi adoptado no passado por muitas empresas de telecomunicações para dimensionarem as suas redes.

O caminho completo de uma ligação telefónica inclui o percurso do sinal sonoro no ar desde a boca do locutor até ao microfone e do altifalante até ao canal auditivo do ouvinte, para além do percurso do sinal eléctrico através de todo o sistema de comunicação. A atenuação total deste caminho constitui o equivalente de referência (*ER*). A título de exemplo apresentar-se-á o significado de alguns equivalentes de referência:

- 0 dB - corresponde a uma pessoa falando a 4 cm do ouvido de quem escuta (voz normal);
- 25 dB - corresponde a dois interlocutores conversando a uma distância de 70 cm (voz normal);
- 36 dB - corresponde a dois interlocutores conversando a uma distância de 3 m (voz normal).

O antigo CCITT (actual ITU-T) recomendava um equivalente de referência máximo entre assinantes de 36 dB, considerando uma ligação na qual estavam envolvidos dois países de tamanho médio. Este equivalente é distribuído do seguinte modo:

[9] Embora muitos dos conceitos apresentados nesta sub-secção tenham sido introduzidos no contexto das redes analógicas ainda é frequente aparecerem hoje em dia na literatura especializada.

- 20.8 dB para a rede nacional do lado do emissor;
- 12.2 dB para a rede nacional do lado do receptor;
- 3 dB para a rede internacional.

Os 3 dB da rede internacional correspondem ao máximo de 6 ligações a 4 fios em cadeia (6×0.5 dB). A distribuição dos equivalentes de referência numa rede de telecomunicações era feita de acordo com um **plano de transmissão**. O plano de transmissão devia definir os valores máximos extremo-a-extremo para os principais factores condicionantes da transmissão (atenuação, ruído, ecos, diafonia, etc.) e indicar a distribuição desses valores pelas diferentes partes constituintes da a rede. Na Figura 1.15 representa-se um exemplo de um plano de transmissão, que surge em consequência directa das recomendações do ITU-T. Note-se que, nesta figura se usa a definição de rede local proposta pelo ITU-T.

Um plano de transmissão apresentava não só os valores dos equivalentes de referência máximos, mas também os mínimos, já que um sinal de voz com volume muito elevado não é desejável.

Nas redes de telecomunicações analógicas os equivalentes de referência podiam variar significativamente de ligação para ligação. Essa não uniformidade no desempenho deve-se, quer às variações da atenuação das linhas de assinante e junção, quer às conversões de dois para quatro fios e vice-versa (como se verá no Capítulo 4, por cada conversão tem-se uma atenuação igual ou superior a 3 dB).

Como já se referiu, será conveniente que as ligações comutadas tenham um desempenho tão uniforme quanto possível, o que nas redes analógicas só era conseguido dentro de certos limites. Para atingir esses objectivos usavam-se sinais de teste, que em todas as troncas eram ajustados para níveis de potência especificados em cada central de comutação, o que era conseguido através da medida do nível desses sinais de teste e da sua comparação com um nível de referência. Como as perdas de transmissão variam com a frequência, esses sinais de teste eram especificados para frequências bem definidas, usualmente 800 ou 1000 Hz para os circuitos de fonia.

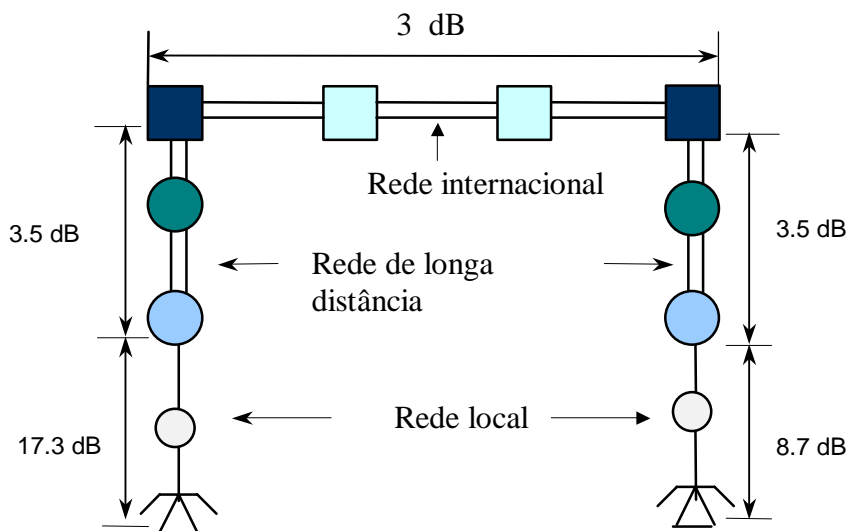



Figura 1.15 Exemplo de um plano de transmissão de uma rede telefónica analógica.

O processo referido levou à definição do conceito de **ponto de nível de transmissão** (PNT). Cada ponto na rede, onde os sinais convencionais de teste eram medidos, designava-se por ponto de nível de transmissão. O nível de transmissão em cada ponto é a relação entre a potência do sinal de teste nesse ponto e a potência do mesmo sinal num ponto de referência arbitrário, designado por **ponto de nível zero de transmissão** (PNT0), e é expresso em dBr. O ponto de nível zero é, assim, caracterizado por 0 dBr. Nas redes telefónicas analógicas o ponto de nível zero era, normalmente, definido como um ponto na entrada de determinados centros de trânsito primários.

A potência de um sinal medida no ponto de nível zero é expressa em dBm0. Se o sinal de teste for igual a 0 dBm0 no ponto de nível zero, então o valor do nível de transmissão é igual à potência real do sinal de teste em cada PNT.

O facto da qualidade de uma ligação telefónica analógica depender do percurso seguido deve-se ao facto de nas redes analógicas os factores perturbadores tais como perdas, ruído e distorção se acumularem ao longo do percurso. Este problema é resolvido quando se usam técnicas digitais, devido ao facto dos sinais serem regenerados nos repetidores em alternativa à simples amplificação do caso analógico. Assim, nas redes que usam transmissão e comutação digitais é possível obter-se um desempenho praticamente uniforme, como é o caso das redes digitais integradas, ou das redes digitais com integração de serviços que são usadas nos dias de hoje.

1.5.1.3 Rede Digital Integrada

Uma **Rede Digital Integrada** (RDI) é definida como sendo uma rede na qual todas as centrais de comutação são digitais e o tráfego nas junções e nas troncas é transportado em sistemas de transmissão digital. Além disso, a sinalização entre as centrais, que como se viu é da responsabilidade do plano de controlo, é assumida como sendo do tipo canal-comum, como é o caso do Sistema de Sinalização nº7.

Dentro da RDI todos os canais de tráfego são em formato digital (PCM), sendo, por conseguinte, a conversão analógico-digital requerida somente nas suas fronteiras, que geralmente se situam na entrada das centrais locais (ver Figura 1.16). O passo seguinte de evolução consiste em proporcionar transmissão digital até ao utilizador e neste caso, somos levados às **Redes Digitais com Integração de Serviços** (RDIS), que serão analisadas posteriormente. Outro aspecto, digno de realce nestas redes é que a comutação digital é sempre feita a quatro fios, de modo que, todos os circuitos dentro da RDI são circuitos a quatro fios. A qualidade de transmissão de uma RDI apresenta os seguintes atributos:

- As perdas de transmissão são independentes do número de troços e centrais presentes numa ligação;
- As ligações apresentam um nível mais baixo de ruído do que as analógicas correspondentes;
- As ligações são mais estáveis do que nas redes analógicas a dois fios.

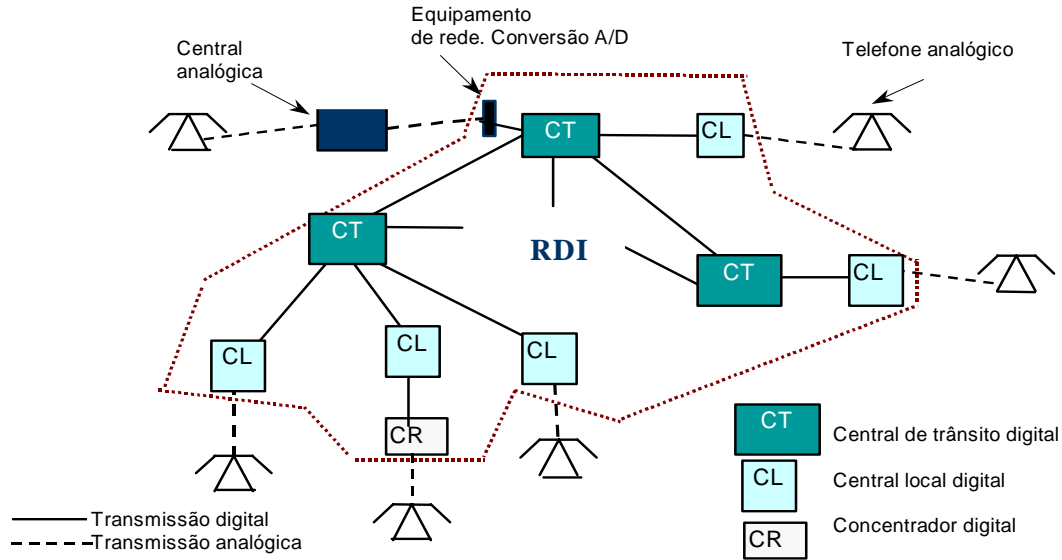



Figura 1.16 Definição de uma rede digital integrada. As fronteiras da RDI são delimitadas a ponteado.

O primeiro atributo é particularmente importante para um operador de telecomunicações, pois garante que as perdas de transmissão nas fronteiras da RDI se mantêm constantes para todos os tipos de ligações, ou seja, é possível garantir uniformidade no desempenho, o que era difícil de satisfazer no caso analógico. As perdas de transmissão são agora escolhidas de modo a garantir um equivalente de referência apropriado e simultaneamente satisfazer os requisitos impostos pela estabilidade e ecos, requisitos estes que analisaremos em detalhe no Capítulo 4. Um valor típico para essas perdas é, por exemplo, 6 dB.

1.5.2 Rede Digital com Integração de Serviços

A Rede Digital com Integração de Serviços (RDIS)[10] resulta da evolução natural da rede telefónica. A rede telefónica foi projectada simplesmente para tráfego de voz sobre linhas analógicas, mas na década de cinquenta foi introduzido o modem para transportar dados sobre essa infra-estrutura.

Contudo, devido às limitações das velocidades de transmissão e qualidade dos modems, os operadores de telecomunicações criaram uma rede digital alternativa à rede de voz, para suportar a transmissão de dados com maior velocidade e melhor qualidade, a rede pública de dados, analisada em traços muito gerais na secção seguinte.

A RDIS surge como tentativa de integrar todas as redes públicas (telefónica, dados, etc.) numa única rede, com um **acesso único ao assinante**. Assim, o utilizador pode através de uma única linha de assinante ter acesso a uma grande diversidade de serviços, como voz, dados, imagem, texto, etc., com uma característica fundamental, que é a de todos serem digitais.

[10] Em notação anglo-saxónica ISDN (*Integrated Services Digital Network).*

A evolução para a RDIS só é possível com a **digitalização da linha de assinante** (ou lacete de assinante[11]), o que permitiu eliminar o fosso analógico existente nas redes RDI. Na linha digital de assinante continua-se a usar a linha telefónica a 2 fios (pares simétricos), requerendo, contudo, um grande esforço de processamento de sinal, para garantir nessas linhas uma transmissão digital com qualidade.

Uma rede RDIS tem possibilidade de oferecer dois tipos de acessos: acesso básico e acesso primário. O primeiro tipo coloca à disposição do utilizador dois canais para transmissão de voz a 64 kb/s e um canal para dados a 16 kb/s, totalizando um débito de 144 kb/s. O acesso básico por sua vez disponibiliza 30 canais de voz com o mesmo débito e um canal de dados a 64 kb/s perfazendo cerca de 2 Mb/s.

Como os débitos oferecidos pelo RDIS são relativamente modestos a ITU-T avançou com o conceito de **RDIS de banda larga** e publicou uma série de normas no sentido de dar substância a esse conceito. Com esta evolução a RDIS passaria a ter também capacidade para suportar serviços de vídeo e de transmissão de dados a alta velocidade, para além dos serviços RDIS tradicionais e o acesso do utilizador à rede seria efectivado a débitos de várias dezenas de Mb/s. A implementação do RDIS de banda larga implicava a implementação de um novo paradigma de transferência de informação, designado por ATM (*Asynchronous Transfer Mode*). Fundamentalmente, o ATM é um protocolo de comutação rápida, que foi concebido no sentido do mesmo comutador ter capacidade para comutar todos os tipos de serviço oferecidos pela rede. Esta capacidade de integrar a comutação de todos os serviços num único elemento de rede representou um salto muito significativo relativamente à filosofia de comutação subjacente ao RDIS tradicional, onde é necessário ter um comutador digital de circuitos para os sinais de voz e um comutador de pacotes para o tráfego de dados. Devido aos elevados custos, o RDIS de banda larga nunca viu a luz do dia, mas ficou a tecnologia ATM, que foi adoptada pela indústria de telecomunicações como uma boa solução para a comutação de dados a muito alta velocidade.

1.5.3 Rede de dados pública

Uma rede de dados é uma rede que permite a troca de informação digital entre computadores, terminais e outros dispositivos processadores de informação, usando diferentes ligações e nós. Como já foi referido a rede de dados pode-se segmentar em três grupos: LAN que é uma rede localizada numa área geográfica limitada (edifício ou campus) e geralmente pertencente a uma única organização; MAN é uma rede cujos pontos de acesso se localizam numa área metropolitana; WAN pode estender-se por várias cidades e mesmo países.

A rede telefónica, cujos traços gerais já foram analisados, não é apropriada para a transmissão interactiva de dados, pois esta é projectada para fornecer serviços com maior duração e com frequências de pedidos de acesso à rede menor. Nem os elementos de controlo nos comutadores, nem a capacidade dos canais são capazes de acomodar pedidos com muita frequência para mensagens muito curtas.

Surgiu, assim, a necessidade de projectar uma rede com uma filosofia de operação claramente distinta da das redes telefónicas. A tecnologia de encaminhamento base escolhida para essas redes foi a **comutação de pacotes**.

[11] Também designado por lacete local.

A Figura 1.17 ilustra o funcionamento de uma rede baseada na comutação de pacotes. Cada mensagem na fonte é dividida em pequenas unidades designadas por **pacotes**, para transmissão através da rede. Esses pacotes, também designados por **datagramas**, para além da informação propriamente dita incluem um cabeçalho, com informação do endereço do destinatário, da fonte, o número do próprio datagrama e outra informação de controlo.

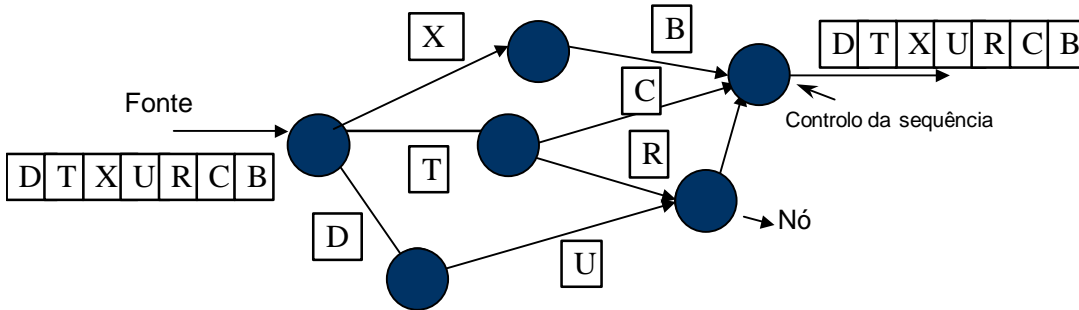



Figura 1.17 Princípio de operação de uma rede de comutação de pacotes.

Os datagramas pertencentes a uma determinada mensagem são enviados pela rede independentemente, podendo seguir percursos diferentes até ao seu destino, onde são agregados de modo a originar a mensagem inicial. Neste tipo de comunicação não é necessário estabelecer uma ligação prévia com o destinatário, pois o cabeçalho contém o endereço do destino final e cada nó, através da leitura desse cabeçalho, está em condições de definir o trajecto a seguir. Este tipo de ligação designa-se por *connectionless*. Este tipo de rede permite garantir um nível de segurança bastante elevado, na medida em que qualquer intruso na rede somente consegue obter fragmentos da mensagem transmitida. Além disso, nesta rede não existe um ponto de falha único, porque se um nó, ou uma ligação falham, ou são sabotados, existem sempre ligações e nós alternativos. Além disso, o controlo deste tipo de rede é distribuído por vários nós, não havendo uma estrutura hierárquica como nas redes telefónicas. Foram estas vantagens que levaram à implementação em 1967 nos Estados Unidos de uma rede de dados baseada neste princípios, designada por ARPANET (*Advanced Research Projects Agency NETwork*), a qual evoluiu para uma rede à escala mundial, ou seja a Internet. O protocolo IP (*Internet Protocol*) é baseado nessa filosofia de interligação fazendo com que as redes IP não estejam em condições de garantir um serviço com uma qualidade pré definida em termos de atraso, erros ou débito. Esses serviços designam-se por isso serviços ao melhor esforço (*best-effort*), indicando que a rede tentará fazer o melhor que pode.

As redes de dados públicas como foram concebidas no início usavam, contudo, um conceito um pouco diferente daquele que foi exposto. Nestas redes, antes de se iniciar a transmissão da mensagem, tinha-se uma fase inicial para estabelecer uma ligação lógica com o destinatário. Assim, o primeiro pacote que é enviado é responsável por estabelecer uma ligação lógica através da rede, designada por **circuito virtual** e todos os pacotes correspondentes à mensagem seguem por essa ligação[12]. Este processo tem algumas semelhanças com aquilo que acontece com a comutação de circuitos nas redes telefónicas,

[12] Os protocolos X.25 e *Frame Relay* usados nas redes públicas são baseados em circuitos virtuais.

mas há uma diferença fundamental: No caso do circuito virtual a ligação não é dedicada, ou seja diferentes circuitos virtuais podem partilhar a mesma ligação. Para que isso seja possível os pacotes recebidos estão sujeitos a um processo de armazenamento/expedição em cada nó.

Neste caso, o cabeçalho do pacote necessita de conter a identificação do circuito virtual e em cada nó não é necessário tomar decisões sobre o encaminhamento da informação, como acontecia nas ligações com datagramas. Esta simplificação permite às redes com circuitos virtuais escoar tráfego com débitos mais elevados e com maior rapidez do que as redes com datagramas. Perde-se, no entanto, a segurança, flexibilidade e fiabilidade associadas à tecnologia dos datagramas.

Em síntese, pode dizer-se que a comutação baseada em circuitos virtuais é mais adequada para transmissões longas e com débitos elevados, enquanto a comutação com datagramas é preferível para transmissão de dados de curta duração.

O paradigma de comutação ATM também é baseado em circuitos virtuais. A diferença essencial é que na comutação de pacotes os pacotes têm dimensão variável, enquanto no ATM têm valor fixo e uma dimensão muito inferior. Para vincar essa diferença, nas redes ATM usa-se a designação de **célula** em vez de pacote. Uma célula é constituída por 53 octetos, sendo 5 usados para cabeçalho e os restantes para informação.

A importância cada vez maior do protocolo IP associada à massificação da Internet e a necessidade de usar este protocolo para suportar serviços diferentes do serviço de dados, como por exemplo a voz e o vídeo levou ao desenvolvimento de estratégias apropriada para também ser possível garantir qualidades de serviço pré determinadas nas redes IP. O protocolo MPLS (*Multiprotocol Label Switching*) vem exactamente nesse sentido. A base deste protocolo consiste em canalizar todos os pacotes com o mesmo destino através de uma espécie de túnel virtual associando-lhe uma **etiqueta** (*labe*l). O encaminhamemto nas redes MPLS é feito unicamente através da identificação da etiqueta, permitindo aumentar significativamente a rapidez de comutação e consequentemente reduzir o atraso do pacote na rede.

1.5.4 Redes híbridas fibra-coaxial

As redes de distribuição de televisão por cabo ou CATV (*CAble* TV) são caracterizadas por usarem uma infraestrutura em fibra óptica para servir núcleos (células) de algumas centenas de utilizadores (200 a 1000), seguida de uma rede em cabo coaxial até às instalações do utilizador (ver Figura 1.19). Por essa razão são designadas por **redes híbridas fibra-coaxial**, embora muitas vezes apareçam referidas na imprensa não especializada simplesmente pela designação de **redes de cabo.** Estas redes foram inicialmente projectadas para oferecer serviços distributivos analógicos (televisão), tendo evoluído posteriormente de modo a oferecerem também serviços distributivos digitais e serviços interactivos como é o caso do acesso à Internet. Na sua componente distributiva o servidor situado na cabeça da rede distribui através da rede vários canais de televisão usando multiplexagem por divisão na frequência. Cada utilizador tem por essa razão acesso a todos os canais e escolhe o canal desejado através da simples sintonização do

televisor.[13] No caso da televisão analógica cada canal da televisão vai modular uma portadora de radiofrequência usando modulação AM-VSB e no caso da vertente digital usa-se modulação PSK ou QAM.

A atribuição espectral típica dos diferentes serviços numa rede híbrida fibra-coaxial está representada na Figura 1.18. Os canais de televisão fazem uso da chamada banda directa situada entre os 111 e 750 MHz, embora essa banda nas redes mais modernas possa ir até 1 GHz. A parte superior da banda (entre os 550 e 750 MHz) é reservada para os canais de televisão digitais, ou como canal descendente, entre a cabeça da rede e o utilizador, para os serviços interactivos. A banda de radiofrequência típica reservada para cada canal analógico é de 8 MHz, sendo que a largura de banda de vídeo nominal é de 5 MHz. A via de retorno, situada entre os 5 e 65 MHz, funciona como canal ascendente para comunicação entre os clientes e a cabeça da rede e desempenha um papel importante na implementação dos serviços interactivos.

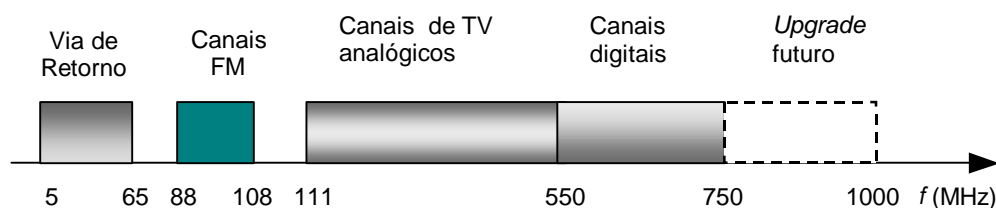



Figura 1.18 Espectro de radiofrequência típico de uma rede híbrida.

A existência de uma banda de retorno com uma gama deveras limitada (60 MHz) é uma das principais limitações das redes híbridas e irá ser um condicionante importante na utilização dessas redes como plataformas de acesso de banda larga no futuro. Como o meio é partilhado, o número de utilizadores activos em cada célula vai influenciar directamente o débito que cada um deles pode usufruir. Para conseguir aumentos significativos nesses débitos será necessário reduzir o número de assinantes por célula, o que implica aproximar a fibra óptica do utilizador. Essa estratégia irá conduzir à eliminação de toda a componente de amplificação de rádio-frequência[14] dessas redes e à transformação da rede coaxial numa rede totalmente passiva.

Para analisar com mais detalhe a estrutura de uma rede híbrida concentremo-nos na Figura 1.19. Nessa figura a ligação entre a cabeça de rede e o nó de acesso é realizada por uma simples fibra óptica. Numa rede real essa ligação é muito mais complexa e pode conter vários níveis hierárquicos, sendo realizada pela rede da camada de transporte. A parte coaxial (rede coaxial) corresponde à componente de acesso da rede híbrida. Essa componente inicia-se no nó de acesso óptico onde tem lugar a conversão do sinal do domínio óptico para o domínio eléctrico e no caso das redes com capacidade para suportar serviços interactivos (bidireccionais) também do domínio eléctrico para o óptico.

---

[13] Note-se que nas redes de distribuição de vídeo sobre IP o utilizador só tem acesso em cada instante a um único canal. Quando pretende vizualizar outro canal tem de o solicitar enviando um sinal de controlo para a cabeça de rede.

[14] Será de notar que esses amplificadores são uma importante fonte de ruído e de origem de falhas e por isso
a sua eliminação vai contribuir para aumentar a qualidade do sinal e a fiabilidade da rede.

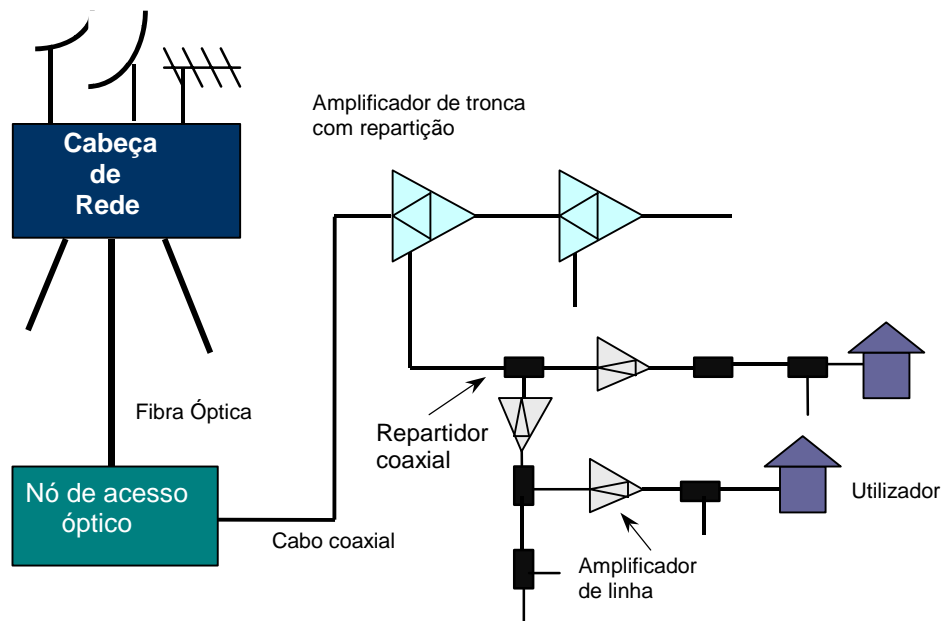



Figura 1.19 Estrutura de uma rede híbrida fibra coaxial.

A rede coaxial tem uma topologia física em árvore. A repartição dos sinais de radiofrequência pode ser feita através dos amplificadores de tronca ou a partir de repartidores passivos. Para além desse amplificadores tem-se ainda os amplificadores de linha que são usados para compensar a atenuação do cabo coaxial e dos repartidores passivos. Será de notar que ambos os tipos de amplificadores terão de ser bidireccionais, para assegurar serviços interactivos nessas redes (para mais detalhes consultar referência [Jo99]).

1.5.5 Redes celulares

O conceito básico subjacente às comunicações celulares consiste em dividir as regiões densamente povoadas em várias regiões de pequena dimensão, designadas por **células.** Cada célula tem uma estação base que proporciona uma cobertura via rádio a toda a célula. Como se mostra na Figura 1.20 cada estação base está ligada a uma central de comutação de móveis, designada por MSC (*Mobile Switching Centre*).

Os componentes básicos da rede são, assim, os telefones móveis, as estações de base e os MSC. Cada MSC controla todas as chamadas móveis entre as células de uma determinada área e a central local.

A estação de base está equipada para transmitir, receber e encaminhar as chamadas para, ou de, qualquer unidade móvel dentro da célula para o MSC. A célula compreende uma área reduzida (geralmente poucos quilómetros quadrados), o que permite reduzir a potência emitida pela estação de base até um nível em que a interferência nas células vizinhas é negligenciável. Tal permite que a mesma radiofrequência seja usada para diferentes conversações em diferentes células, sem existir o perigo de interferência mútua.

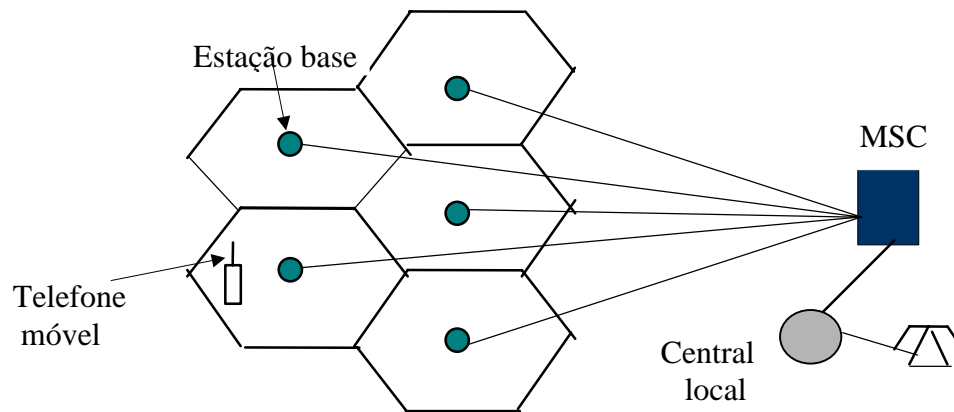



Figura 1.20 Estrutura básica de uma rede celular.

1.5.5.1 Operação

Cada célula tem disponíveís vários rádiocanais para tráfego de voz e um, ou mais, para sinalização de controlo. Quando o telefone móvel é ligado, o seu microprocessador analisa o nível de sinal dos diferentes canais de controlo pertencentes a uma mesma MSC, e sintoniza o seu receptor para o canal com o nível mais elevado. Esta operação designa-se por **auto-localização**, e permite estabelecer um lacete entre a unidade móvel e a estação de base, que será mantido enquanto o telefone estiver ligado. Periodicamente, o nível de sinal dos diferentes canais de controlo continua a ser analisado, garantindo-se, assim, o estabelecimento de lacetes com outras estações base, na eventualidade de a unidade móvel se deslocar para outras células. Outra operação associada ao estabelecimento de uma ligação é o **registo de presença.** No início da ligação e posteriormente, em intervalos regulares, o telefone móvel envia informação da sua presença para a MSC mais próxima. Essa informação é armazenada numa base de dados e permite à MSC ter uma ideia aproximada da localização do móvel.

Quando a unidade móvel pretende realizar uma chamada, transmite o número do destinatário para a estação base, usando o canal de controlo. A estação base envia então essa informação para o MSC, juntamente com seu número de identificação. Imediatamente, o MSC atribui um radiocanal de voz bidireccional para o estabelecimento da ligação entre o telefone móvel e a estação base. Depois de receber esta informação, o microprocessador do telefone móvel ajusta o sintetizador de frequência para emitir e receber nas frequências atribuídas. Logo que o MSC detecta a presença da portadora da unidade móvel no canal desejado, a chamada ou é processada pelo próprio MSC, ou é enviada para a central local para aí ser processada.

Uma função importante da MSC consiste em localizar o destinatário, no caso em que este é um móvel. A função de localização está associada ao ***paging.*** Depois de localizado, o sinal de chamada pode em seguida ser ouvido no telefone móvel do destinatário.

Quando a estação base de uma determinada célula detecta que a potência do sinal emitido por uma determinada unidade móvel desce abaixo de um determinado nível, sugere à MSC para atribuir o comando dessa unidade a outra estação base. A MSC, para localizar o móvel, pede às celulas vizinhas informação sobre a potência do sinal por ele emitido, sendo atribuído o comando do móvel à estação base que reportar um nível de

sinal mais elevado. Um novo canal de voz é atribuído a essa unidade móvel pelo MSC, sendo a chamada transferida automaticamente para esse novo canal. Este processo designa-se por ***handover*** e dura cerca de 200 ms, o que não é suficiente para afectar uma comunicação de voz.

1.6 Redes do Século XXI

Não é possível definir com rigor como vão ser as redes de telecomunicações que vão ser construídas no futuro. Há contudo uma certeza: estas redes vão ser concebidas num ambiente e num tempo em que as características dos serviços e do tráfego são completamente diferentes daquelas que serviram de base à concepção das redes telefónicas públicas comutadas, pois como se sabe essas redes foram concebidas para dar resposta ao tráfego de voz e hoje o tráfego dominante é o tráfego de dados.

Esse novo cenário impõe novas exigências que podem ser traduzidas em duas palavraschave: banda larga e convergência. A banda larga implica que a rede tenha capacidade para disponibilizar canais de comunicação entre os utilizadores com débitos superiores aos permitidos, nomeadamente, pelo acesso básico das redes RDIS. Embora a fronteira que delimita o conceito de banda estreita e banda larga não esteja pré definido, e seja variável ao longo do tempo, hoje há algum consenso em considerar como acesso de banda larga só aquele acesso que permita débitos superiores a 2 Mbit/s. O enfoque nos serviços de banda larga irá ter certamente reflexo nas tecnologias e no suporte da transmissão usada no acesso: os pares simétricos (cobre) e o cabo coaxial irão perdendo progressivamente importância e o seu lugar será ocupado pela fibra óptica.

A convergência poderá ser perspectivada quer em termos de serviços, quer em termos de redes. O último tipo de convergência tem como objectivo reduzir o número de tecnologias usadas, simplificar a operação das redes e por conseguinte reduzir o seu custo de exploração. A convergência deverá ter lugar, quer a nível da camada de rede de serviço, quer a nível da camada de rede de transporte. Esse movimento ao nível da primeira camada poderá levar no início ao colapso da rede telefónica pública comutada e da rede de dados pública numa única rede, e numa etapa seguinte é possível que esse movimento englobe também a rede de distribuição de televisão por cabo.

Um bom exemplo é o caso do programa "BT 21st *Century Network*" levado a cabo pela British Telecom no Reino Unido. Esse programa orçamentado em 10 mil milhões de libras (quinze mil milhões de euros), iniciado em 2004 e com a duração de cinco anos tem por objectivo mudar toda a infraestrutura da rede da BT de modo a adaptá-la aos novos paradigmas. A rede de serviço do 21st *Century Network* (ver Figura 1.21) é baseada na tecnologia IP/MPLS e resulta da convergência de quatro redes: PSTN (*public switched telephone network*), DPCN (*data public communication network)*, ATM, e IP. Por sua vez a rede de transporte é baseada na tecnologia OTN (*optical transport network*), que surge como fruto da convergência entre as tecnologias PDH[15], SDH e WDM. Um equipamento fundamental nessa rede é a plataforma de acesso multiserviço. Essa plataforma actua como interface entre o mundo IP/MPLS e os diferentes serviços. Por exemplo, no caso do serviço telefónico funciona como central local, no caso dos serviços Ethernet como

[15] A tecnologia PDH (*plesiochronous digital hierarchy)* irá ser estudada no Capítulo 6.

comutador Ethernet, no caso do ADSL como DSLAM[16], etc. Esse programa também coloca muita ênfase sobre a utilização da fibra óptica no acesso. Essa utilização poderá conduzir à substituição total do cobre no caso das novas construções, ou à sua substituição parcial no caso das construções já existentes.

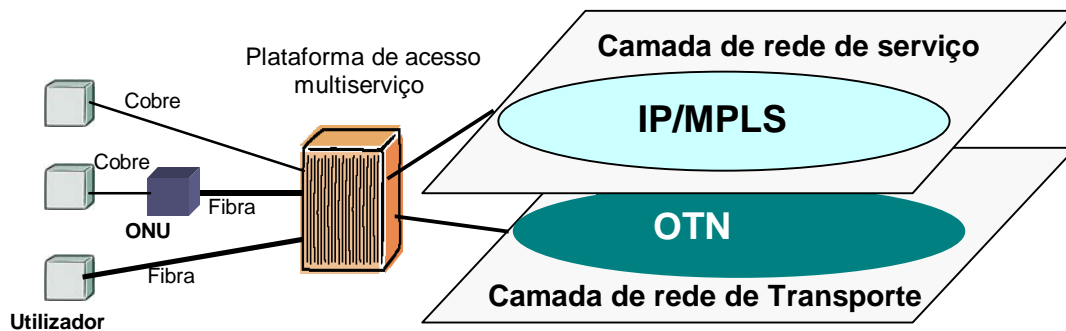



Figura 1.21 Visão esquemática da rede BT 21st *Century Network*.

Este movimento iniciado por uma operadora que tem sido pioneira na inovação tecnológica irá certamente em breve estender-se a outros países incluindo Portugal. É claro que as soluções tecnológicas a adoptar poderão não ser exactamente as mesmas. Por exemplo, a tecnologia SDH da nova geração, também poderá ter um papel importante na concepção das futuras camadas de transporte, sempre que se requererem níveis de capilaridade[17] maiores e uma dinâmica mais rápida do que aqueles que são permitidos pelas tecnologia OTN[18]. Fica ainda uma outra certeza: essas mudanças requerem investimentos colossais, enormes corpos de engenharia, e quando se iniciarem em Portugal colocarão desafios e ao mesmo tempo apresentarão oportunidades como a engenharia das telecomunicações portuguesa nunca esteve sujeita.

1.7 Problemas

1.1 Qual é a principal limitação das topologias em barramento e em árvore? Como é que essa limitação pode ser ultrapassada?

1.2 Explique se o plano de controlo intervém no processo de estabelecimento de uma chamada telefónica.

1.3 Indique quais são as diferenças essências entre uma rede de transporte e uma rede de serviços. Dê exemplos de redes de serviço.

---

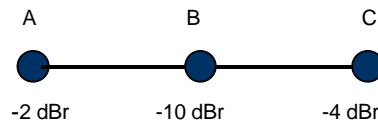
[16] As funcionalidades do DSLAM (*digital subscriber line access multiplexer*) serão estudadas no Capítulo 5.
[17] As redes OTN só permitem caminhos com débitos iguais ou superiores a 2.5 Gb/s. Pode haver interesse para alguns serviços débitos inferiores a estes (maior capilaridade).
[18] Para detalhes sobre o OTN ver as normas do ITU-T G.709, G.872 e G.959.

1.4 Admita que a rede de serviço representada na Figura 1.9 requer que a conectividade proporcionada pelo transporte conduza a uma topologia lógica em estrela dupla. Represente num esquema o diagrama de conectividade (conjunto de caminhos) da rede de transporte que garantem essa topologia lógica. Qual é a vantagem que vê na utilização de uma topologia lógica em estrela dupla em detrimento da estrela simples.

1.5 Admita que a rede de transporte de uma rede híbrida fibra-coaxial tem uma topologia física em anel com quatro nós. Tendo presente esse dado faça uma representação esquemática de toda a rede híbrida (camada de serviço+transporte).

1.6 Na sua perspectiva, quais foram as principais modificações que foi necessário fazer nas redes híbridas, para que estas pudessem evoluir de redes com capacidades meramente distributivas (televisão) para redes capazes de proporcionaram serviços interactivos, como é o caso do acesso à Internet.

1.7 Considere os seguintes valores de potência de um sinal: 50 μW, 1 mW e 100 mW. Exprima essas potências em dBm e dBW. Calcule em dBV e dBμV as tensões que essas potências produzem numa resistência de 600 Ω e 50 Ω.

1.8 Com base na figura representada abaixo determine: a) A potência do sinal medida no ponto B, admitindo que no ponto de nível zero de transmissão se injecta uma potência de 1 mW; b) O valor do ganho (perdas) que o sinal sofre quando se propaga de A a C; c) O valor da potência de ruído medida no ponto de nível zero e em C, admitindo que o nível absoluto da potência em B é de –60 dBmp.

1.9 Considere que na componente coaxial de uma rede híbrida se têm dois amplificadores ligados por um troço de cabo coaxial com o comprimento de 80 m. Tenha presente que a atenuação do cabo coaxial (@ 750 MHz) é de 8 dB/100m, que para a mesma frequência os amplificadores têm as seguintes especificações: $F1$=5 dB, $G1$=15 dB, $F2$=7 dB e $G2$=35 dB, e que a largura de banda equivalente de ruído do sinal de televisão é de 4.75 MHz. Se a temperatura equivalente de ruído na entrada do primeiro amplificador for a temperatura padrão de 290 K, determine a potência de sinal na entrada para garantir uma relação sinal-ruído de 30 dB à saída, admitindo que o canal em análise é transmitido na frequência de 750 MHz. Exprima essa potência em pW, dBW e dBm.

1.8 Referências

[Ba66] P. Baran, “On distributed communications networks”, *IEEE Trans. Commun. Syst.*, vol. CS-12, pág. 1-9, Março 1964.

[Br95] J. Bray, *The Communications Miracle*, Plenum Press, 1995.

[Fl95] J .E. Flood, *Telecomunications Switching, Trafic and Networks*, 1ª edição, Prentice Hall International, 1995.

[Gu03] A. Gumaste, T. Antony, *DWDDM network designs and engineering solutions,* Cisco Press, 2003.

[Jo99] João J. O. Pires, Sistemas de Comunicação Óptica, IST, 1999.

[Ka66] K. C. Kao e G. A. Hockman, “Dielectric fiber surface waveguides for optical frequencies,” *Proc. IEE*, 133, pág. 1151-1158, Julho 1966.

[Mc90] J. C. McDonald, *Fundamentals of Digital Switching,* 2ª edição, Plenum Press, 1990.

[Nu92] M. S. Nunes e A. J. Casaca, *Redes Digitais com Integração de Serviços*, Editorial Presença, 1992.

[Ra94] T. Ramateke, *Networks*, Prentice-Hall International, Inc., 1994.

[Sa99] R. Santos, *Olhos de Boneca*, Edições Colibri/Portugal Telecom, 1999

[Sh99] W. A. Shay, *Understanding data communications & networks*, 2ª edição, Brooks/Cole Publishing Company, 1999.

2 Serviços em Telecomunicações

2.1 Tipos de serviços e suas exigências

O objectivo de uma rede de telecomunicações é fornecer serviços tais como voz, vídeo, e dados aos clientes. Em geral, esses clientes estão sobretudo interessados no tipo, na qualidade e no custo dos serviços (incluindo a tarifa da rede e os custo dos terminais), sendo-lhes indiferente a arquitectura da rede que os suporta. No entanto, os serviços e as redes constituem um todo, não sendo concebíveis os primeiros sem as segundas e vice-versa.

As redes de telecomunicações convencionais foram moldadas pelo crescimento dos serviços telefónicos ao longo de várias décadas. No entanto, a expansão da Internet o aparecimento de novas tecnologias como a xDSL[19] e novos conceitos como o VoIP[20] ou Vídeo sobre IP, e o aparecimento de uma infra-estrutura de transmissão com uma largura de banda quase ilimitada (fibra óptica) vieram criar as condições para o aparecimento de uma miríade de novos serviços, tal que hoje em dia ainda não é fácil visualizar com nitidez a totalidade dos seus contornos.

Os serviços fornecidos por uma rede de telecomunicações podem ser classificados segundo diferentes classes que estão associadas os diferentes parâmetros: direcção do fluxo de informação, modo de representação da informação, requisitos impostos à rede, etc. Tendo em conta a direcção do fluxo de informação os serviços podem ser:

- **Serviços distributivos**: Serviços caracterizados pela fluxo de informação ser unidireccional de um ponto central da rede para múltiplos utilizadores. Este tipo de serviços ainda pode ser oferecido num ambiente em que o utilizador não tem qualquer controlo sobre o início e a ordem da apresentação (sem controlo da apresentação) como é o caso de difusão de televisão, ou num ambiente em que o utilizador pode influenciar essa apresentação (com controlo de apresentação) como é o caso do vídeo a pedido ou do teletexto.

- **Serviços interactivos:** Serviços caracterizados pelo fluxo de informação ser, normalmente, bidireccional. Estes serviços ainda se podem apresentar segundo várias envolventes: **serviços de conversação, serviços de consulta** e **serviços de mensagem**. O primeiro tipo requer uma transferência de informação extremo-a-extremo em **tempo real** como é, por exemplo, o caso da telefonia ou da vídeo-conferência Os serviços de consulta permitem que o utilizador consulte informação armazenada noutros pontos da rede, como seja o caso de acesso a documentos ou vídeo. Finalmente, nos serviços de mensagem, que têm a particularidade de não operarem em tempo real, e de a troca de informação ocorrer entre entidades com capacidade de armazenamento ou caixas de correio electrónico. As

[19] Qualquer tipo de tecnologia DSL
[20] *Voice over Internet Protocol*

mensagens de correio electrónico ou as mensagens telefónicas são exemplos de serviços de mensagem.

Se a perspectiva de análise for o modo de representação da informação os serviços podem ser divididos em **monomédia** e **multimédia.** Na primeira classe incluem-se os serviços que usam um único modo para representar a informação, como sejam por exemplos os serviços de telefonia, telecópia e televisão, enquanto a última classe inclui os serviços que fazem uso de diferentes modos (texto, gráficos, áudio, imagens e vídeo). Os serviços proporcionados pela Internet são um bom exemplo de serviços multimédia.

Se por sua vez a perspectiva de análise for os requisitos impostos à rede há diferentes parâmetros que podem ser equacionados: **débito binário**, **simetria**, **tipo de ligação,** etc. O débito é o parâmetro que quantifica o valor instantâneo de fluxo de informação. Este débito pode ser **constante** ou **variável**. Os serviços de voz e vídeo digitais são exemplos de serviços de débito constante. Já a voz sobre IP, o vídeo sobre IP, assim como a generalidade dos serviços suportados no acesso à Internet são serviços de débito variável, onde o débito binário pode variar entre limites definidos. O débito de 2 Mbit/s é usado como fronteira de diferenciação de duas classes de serviços: **banda estreita** e **banda larga.** A classe de banda estreita corresponde aos serviços que requerem um débito inferior ou igual a esse valor de fronteira e inclui, por exemplo, o serviço telefónico analógico, RDIS, acesso normal à Internet e comunicações de móveis. A classe de banda larga requer valores superiores a 2 Mbit/s e inclui serviços tais como distribuição de televisão digital, acesso à Internet de alta velocidade, videoconferência, interligação de centros de computação, tele-medicina, etc. A simetria traduz a relação entre as larguras de banda necessárias para os dois sentidos da ligação: sentido ascendente (utilizador-rede) e sentido descendente (rede-utilizador). A classe simétrica requer uma largura de banda idêntica, e a classe assimétrica diferente. Como exemplo da primeira classe pode-se apresentar a telefonia e a videoconferência e como exemplo da segunda o acesso à Internet.

O tipo de ligação exigido permite diferenciar os serviços como sendo em modo não orientado à ligação (*connectionless)* e em modo orientado à ligação (*connection-oriented*). Os serviços que permitem a transferência de informação entre os utilizadores sem necessidade de estabelecimento prévio de ligação são do primeiro tipo, enquanto os segundos requerem esse estabelecimento prévio, recorrendo para isso à intervenção do plano de controlo. Os serviços de voz e vídeo são do tipo orientado à ligação e os serviços de dados podem ser não orientados à ligação (datagramas) caso da Internet, ou orientados à ligação (circuitos virtuais).

Outra especificidade dos serviços é a sua maior ou menor sensibilidade a certos parâmetros directamente associados às características da rede, tais como o atraso e os erros de transmissão. Nos extremos do espectro encontram-se os serviços de voz, que são muito sensíveis a atrasos e tolerantes a erros, e os serviços de dados que são insensíveis a atrasos mas muito sensíveis a erros. No caso específico do serviço de voz (conversação) os tempos de atraso não podem ultrapassar algumas centenas de milissegundos, sob pena de não ser possível manter uma conversação inteligível entre dois interlocutores. Note-se que, esta enorme sensibilidade deste serviço ao atraso, tem sido uma das principais barreiras presenciadas pelo serviço de voz sobre IP e que na realidade tem tido alguma dificuldade em ultrapassar. O tempo de atraso tem duas componentes: o atraso de propagação do sinal e o atraso dependente do equipamento responsável pelo

processamento de sinal. Os atrasos de propagação dependem da distância e da velocidade de propagação no meio de suporte e são independentes do débito binário, no caso em que o suporte de transmissão é não distorcivo. Por sua vez, os atrasos associados ao equipamento são devidos à necessidade de memorização da informação em operações como codificação de fonte, correcção de erros, multiplexagem, e segmentação em pacotes, etc. e são inversamente proporcionais ao débito binário de transmissão.

2.2 Características dos sinais de voz e de vídeo e sua digitalização

2.2.1 Sinais de voz

A voz constitui um vector de informação analógico primordial na comunicação humana directa, ou indirecta, através de uma rede de telecomunicações. O estudo da voz é muito complexo, pois a sua produção depende de um grande número de factores não só de origem fisiológica, mas também de origem psicológica, sendo a sua análise unicamente possível através de uma abordagem estatística.

Os sons associados à voz humana podem ser categorizados em duas grandes classes, dependendo do modo como são produzidos. A primeira classe é constituída pelos sons vozeados (vogais e certas consoantes tais como j, l, m, v, etc.) que são gerados pela vibração das cordas vocais. A segunda classe inclui os sons não vozeados (f, s, p, ch, etc), que são gerados pelo fluxo de ar na boca modulado pelos articulantes (maxilares, língua e lábios). Não será surpresa que essas duas classes de sons apresentem características espectrais muito distintas. Os sons vozeados apresentam uma amplitude elevada e uma variação pseudo-periódica, com um período T, determinado pela vibração das cordas vocais. O seu espectro é discreto com uma frequência fundamental de 100 a 200 Hz nos homens e 200 a 400 Hz nas mulheres (ver Figura 2.1). Os sons não vozeados apresentam em geral amplitudes baixas e uma variação praticamente aleatória. O seu espectro é contínuo e estacionário.

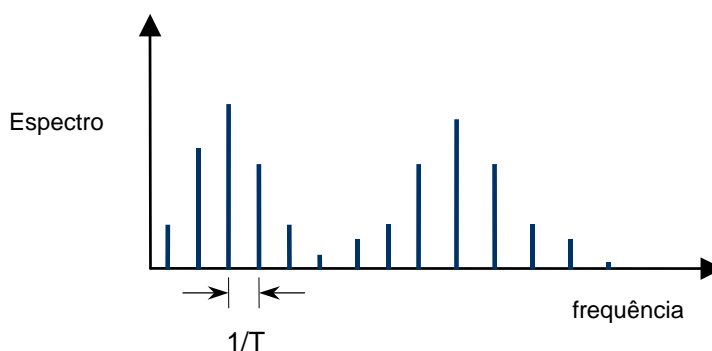



Figura 2.1 Espectro discreto de um som vozeado.

O domínio de frequências dos sinais de voz estende-se de 80 Hz a 12 kHz, com elevada percentagem da riqueza espectral concentrada na banda entre os 150 e 8000 Hz. No entanto, testes subjectivos de inteligibilidade demonstram que a qualidade é, em geral,

satisfatória se se restringir a banda passante a um domínio entre os 300 e os 3400 Hz, embora possam surgir problemas com a distinção de certos sons não vozeados como "f" e "s".

Deste modo, o ITU-T recomenda[21] o uso nos sistemas telefónicos públicos da banda passante entre os 300 e os 3400 Hz, o que traduz um compromisso entre uma transmissão de voz com qualidade e o custo de transmissão de uma larga banda de frequências.

Devido à estrutura temporal irregular da voz (as palavras e as frases são separadas por pausas de duração maior que 100 ms) um sinal de voz activo só está presente no canal de transmissão durante uma determinada fracção do tempo. Medidas efectuadas demonstraram que o sinal de voz só está presente num canal telefónico em aproximadamente 40% do tempo. Assim, essas pausas convenientemente detectadas, podem ser usadas para intercalar outras conversações usando um sistema TASI (*time assignment speech interpolation*), particularmente nas ligações em que o preço dos circuitos é muito elevado, como é o caso das ligações transoceânicas, ou ainda, para transmitir dados juntamente com a voz.

2.2.2 Resposta do ouvido

Outro aspecto importante para a engenharia de telecomunicações é as características da banda audível. Um indivíduo normal com idade compreendida entre os 18 e 25 anos é capaz de detectar tons puros entre 20 Hz e 20 kHz. Porém, com a idade o limite superior de frequência audível reduz-se significativamente (em média um homem de 65 anos tem a 8 kHz, um perda de sensibilidade de 40 dB).

A sensibilidade do ouvido varia com a frequência e com a intensidade sonora. O facto de a sensibilidade do ouvido depender da frequência terá de se reflectir na análise do desempenho das redes telefónicas, particularmente na medida da potência de ruído. Assim, estas medidas são feitas usando um voltímetro de valor eficaz, que incorpora um filtro que simula a variação da sensibilidade do ouvido em função da frequência. Este aparelho designa-se por **psofómetro** e a potência de ruído por ele medida por **potência psofométrica** de ruído, a qual é usualmente expressa em dBmp. Se a potência média de uma fonte de ruído branco for medida numa faixa de 0 a 4000 Hz, depois de passar por um filtro psofométrico, a potência de ruído é reduzida de 3.6 dB relativamente ao caso sem ponderação. No caso da faixa de medida ter sido entre 300 e 3400 Hz (ver Figura 2.2) a diferença passará a ser de 2.5 dB.

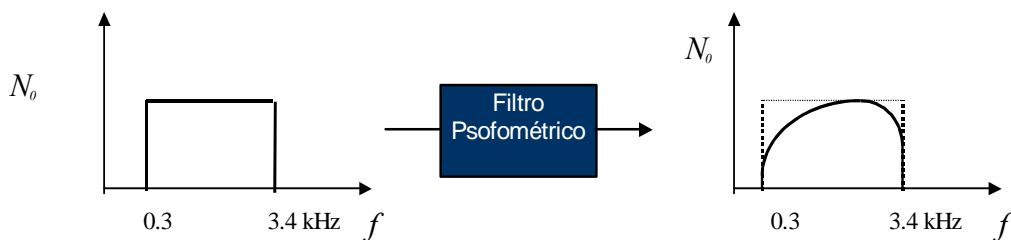



Figura 2.2 Efeito da ponderação psofométrica sobre o ruído branco.

[21] Recomendações G.132 e G.151

Outra característica do ouvido é a sua elevada gama dinâmica, com valores que podem ir acima de 100 dB. Para obter uma boa reprodução da voz é possível, no entanto, utilizar uma gama dinâmica muito menor: 30 dB ou mesmo 20 dB são considerados valores aceitáveis.

A reprodução da música é muito mais exigente do que da voz, requerendo a presença de todas as componentes espectrais perceptíveis pelo ouvido humano. Deste modo, muitos dos sistemas áudio de alta-fidelidade apresentam uma banda entre os 20 Hz e 15 kHz, a qualidade CD requer uma banda entre 20 Hz e 20 kHz e o som de qualidade profissional exige uma banda entre 20 Hz e 40 kHz.

Exemplo 2.1

Para um temperatura ambiente de 290ºK determine a potência de ruído na banda de frequência do canal telefónico, exprimindo o resultado em dBm e dBmp ($k=1.38\times10^{-23}$J/K).

Solução: A potência de ruído na banda B em dBm é dada por

$$N = 10\log_{10}(N_0 B) + 30\,\text{dBm}$$

onde $N_0 = kT$ é a densidade espectral de potência de ruído expressa em Joule e B é a largura de banda do canal telefónico em Hertz. Atendendo aos valores apresentados, a equação anterior conduz a

$$N = -174 + 10\log_{10}(B)\ \text{dBm}$$

Como para o canal telefónico se tem uma banda de B=3.1 kHz, a potência de ruído vale N=-139 dBm. Considerando uma potência psofométrica, o nível de potência está 2.5 dB abaixo, ou seja -141.5 dBmp.

2.2.3 Sinal de vídeo

Como exemplos de sistemas usados para transmitir sinais de vídeo podemos ter a televisão convencional, a televisão interactiva, a videoconferência, o videotelefone, a televisão de alta definição, etc.

Uma imagem é composta por um conjunto finito de elementos de imagem designados por "pixels", cada um dos quais é caracterizado pela sua posição, brilho (luminância) e cor (crominância).

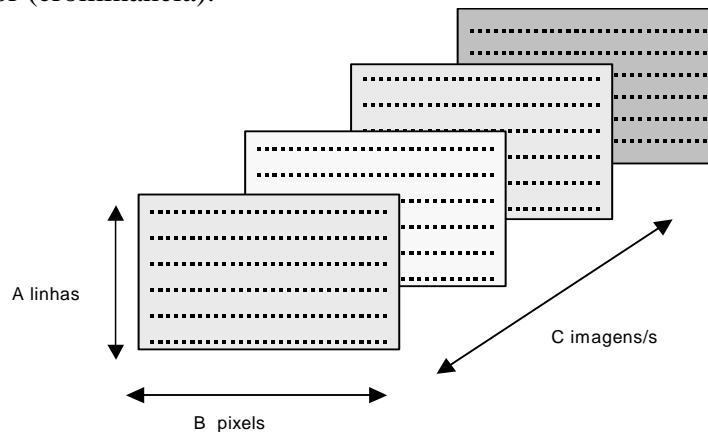



Figura 2.3 Sequência de imagens correspondentes a um sinal de televisão.

Para dar a sensação de movimento contínuo é necessário transmitir mais do que 16-18 imagens/s. Em televisão transmitem-se 25 ou 30 imagens/s. Nos sistemas de televisão convencionais cada imagem para ser transmitida é varrida na câmara de vídeo linha por linha por um feixe de electrões e a sua luminância (brilho) é representada por uma tensão designada por sinal de vídeo. A imagem é reconstruída no tubo de raios catódicos do receptor por um feixe de electrões que varre o ecrã em sincronismo com o feixe da câmara. Este sincronismo é garantido transmitido juntamente com o sinal de televisão um sinal adequado para essas funções.

Deve haver um número suficiente de linhas de modo a proporcionar uma quantidade de detalhe suficiente e a dar uma sensação de continuidade. Diz-se que uma imagem tem baixa definição se tiver menos de 300-400 linhas, média definição entre 500 a 600 linhas e alta definição mais de 1000 linhas. No sistema europeu PAL (*Phase Alternation Line*) o número de linhas transmitido é 625, das quais só 575 são visíveis, sendo as restantes usadas para retornar ao início da imagem.

De acordo com a Figura 2.3 o número de elementos de imagem transmitidos por segundo é dado por $M=ABC$. Considerando C=25 imagens/s, A=575 linhas e para B um valor típico de 720 tem-se $M=10.35\times10^6$ elementos de imagem por segundo. Numa análise muito simplista, cada elemento de imagem pode ser visto como uma amostra da imagem, e assim, recorrendo ao teorema da amostragem somos levados à conclusão que um sinal de televisão requer um largura de banda máxima de $M/2$=5.175 MHz. No entanto, esta análise para além de ser simplista também não entra em consideração com certos parâmetros, como por exemplo o retorno do feixe. Uma análise mais rigorosa, a qual está fora do âmbito desta disciplina, conduz ao valor de 5.5 MHz, para a largura de banda do sinal de televisão. Para a televisão de alta definição tem-se como valores típicos A=1080 linhas e B=1920 elementos de imagem/linha e C=25 imagens/s. A mesma linha de análise conduz a uma largura de banda de 25.92 MHz, ou seja cerca de cinco vezes superior ao valor requerido pela televisão convencional.

2.2.4 Digitalização

Entende-se por digitalização a transformação de um sinal analógico num sinal digital PCM (*Pulse Code Modulation*), enquanto é preservada a quase totalidade do seu conteúdo informativo. Esta operação é realizada por um CODEC (codificador/descodificador), que é responsável pela realização das seguintes operações de processamento do sinal:

- Filtragem;
- Amostragem do sinal analógico;
- Quantificação das amostras discretas no tempo;
- Codificação dos sinais discretos em amplitude e no tempo.

2.2.4.1 Amostragem

Se se pretender reconstituir o sinal original a partir das amostras sem introduzir distorção, o teorema da amostragem diz-nos que o ritmo de amostragem deve ser maior do que o dobro da frequência mais elevada presente no sinal. Já se viu que, na transmissão

telefónica da voz a frequência mais elevada é de 3.4 kHz. Assim, este sinal pode ser amostrado de modo seguro a um ritmo de 8 kHz, sem introdução de distorção adicional. A amostragem do sinal de luminância em televisão é normalmente feita a um ritmo de aproximadamente $f_a = n.f_{sc}$, onde n é um número inteiro e f_{sc} é a frequência da sub-portadora de crominância que é igual a 4.433 MHz. Para um valor típico de n=4, o ritmo de amostragem é de 17.73 MHz. Uma outra alternativa para a digitalização do sinal de televisão consiste em processar em separado os sinais de luminância e de crominância. Neste caso, têm-se as frequência de amostragem de 13.5 MHz e 6.75 MHz, respectivamente, para os sinais de luminância e de crominância.[22]

2.2.4.2 Quantificação

As amostras antes da quantificação podem tomar uma gama infinita de valores e são tão susceptíveis ao ruído como o sinal original. A função da quantificação é aproximar cada uma destas amostras a um determinado nível discreto de amplitude. Obviamente, a operação de quantificação introduz distorção no sinal designada normalmente por **ruído de quantificação.**

A potência deste ruído depende do número de níveis de quantificação usados. Aumentando o número de níveis reduz-se o ruído de quantificação, mas requer-se um maior número de bits para codificar cada nível, o que conduz a um aumento do débito binário. Será, assim, necessário encontrar uma solução de compromisso.

Para a transmissão telefónica de voz, 256 níveis com espaçamento uniforme permitem garantir uma qualidade tal que o ruído de quantificação dificilmente se distingue dos outros ruídos de fundo presentes no canal telefónico. Para garantir na televisão uma qualidade semelhante à de televisão radiodifundida são também usados 256 níveis uniformes.

2.2.4.3 Codificação

No sentido de aproveitar a imunidade dos sinais digitais ao ruído, os níveis de quantificação são normalmente transmitidos usando um código binário, como se representa na Figura 2.5. O número de dígitos binários requeridos para representar cada amostra é igual a

$$N_b = \log_2 L \quad (2.1)$$

onde L representa o número de níveis de quantificação. O débito binário é assim igual a N_b vezes a frequência de amostragem, o que dá para o sinal de voz telefónico 8x8=64 kb/s, considerando que se usam 8 bits por amostra. Para o sinal de televisão usando os dados apresentados ter-se-ia cerca de 108 Mbit/s para o sinal de luminância. Porém, nos sistemas policromáticos para reproduzir a cor é necessário transmitir dois sinais adicionais, cada um com um débito de 54 Mbit/s. Considerando ainda a necessidade de algum "*overhead*" para funções de sincronismo, tem-se para um canal de vídeo policromático um débito de 243 Mbit/s. Outro sinal com interesse é a música.

[22] Recomendação 601 do ITU-T

Usualmente, para gravação usa-se uma frequência de amostragem de 44.1 kHz e uma codificação uniforme com 16 bits/amostra, o que conduz a um ritmo de transmissão da ordem de 0.7 Mb/s.

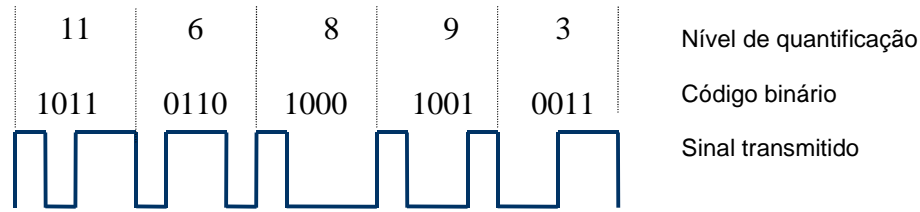



Figura 2.4 Codificação binária.

2.2.4.4 Ruído de quantificação uniforme

No processo de quantificação todas as amostras do sinal pertencentes a um determinado intervalo de quantificação são representadas por um valor discreto correspondente ao centro do intervalo de quantificação.

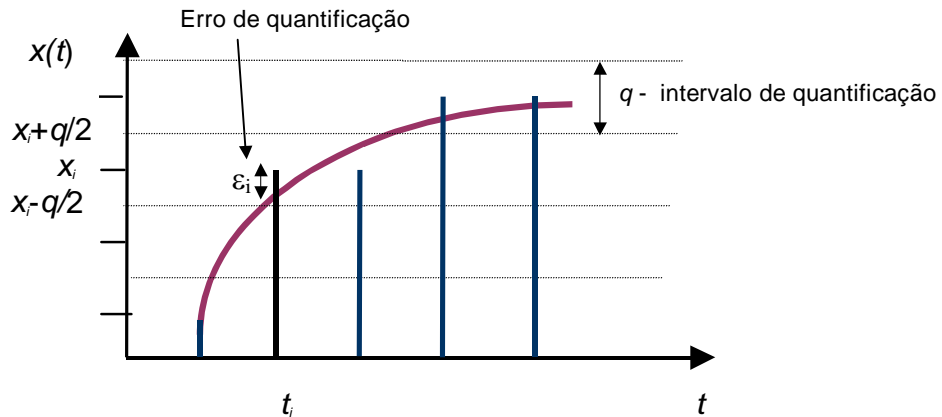



Figura 2.5 Erro de quantificação em quantificação uniforme.

Considere-se uma amostra do sinal $x(t)$ tirada no instante t_i a qual se encontra no intervalo $x_i + q/2 > x(t_i) > x_i - q/2$. Esta amostra irá ser quantificada pelo nível x_i como se mostra na Figura 2.5. Tem-se, assim, um erro de quantificação de $\varepsilon_q = x_i - x$, onde $x = x(t_i)$. Será de notar que:

- O erro ε_q tem uma amplitude limitada a $q/2$, onde q é a largura do intervalo de quantificação;
- Se a amostra $x(t_i)$ tiver uma probabilidade idêntica de se situar em qualquer ponto do intervalo referido a distribuição de ε_q é uniforme, sendo caracterizada pela função densidade de probabilidade

$$p(\varepsilon_q) = \frac{1}{q} \qquad - \frac{q}{2} \le \varepsilon_q \le \frac{q}{2} \tag{2.2}$$

A potência de ruído de quantificação pode, assim, ser calculada através de

$$< \varepsilon_q^2 >= \int_{-\infty}^{\infty} \varepsilon_q^2 p(\varepsilon_q) d\varepsilon_q = \frac{q^2}{12} \tag{2.3}$$

O efeito cumulativo do ruído de quantificação nos sistemas PCM pode ser tratado como ruído aditivo com um efeito subjectivo similar ao do ruído branco. Deste modo, faz sentido falar em relação sinal-ruído de quantificação que se pode definir do seguinte modo:

$$\frac{s}{n_q} = \frac{< x(t)^2 >}{< \varepsilon_q^2 >} \tag{2.4}$$

onde *x(t)* representa o sinal analógico de entrada. O cálculo da potência *s* requer o conhecimento da estatística do sinal $x(t)$. Porém, em muitos casos, a análise do desempenho de um quantificador é baseado em sinais de entrada sinusoidais, já que estes facilitam as medidas e o cálculo da relação sinal-ruído de quantificação. Como um sinal sinusoidal com amplitude A_s tem uma potência igual a $s = A_s^2/2$, usando (2.3) e (2.4) conclui-se que

$$\frac{S}{N_q}(dB) = 10\log_{10}\left(\frac{A_s^2/2}{q^2/12}\right) = 7.78 + 20\log_{10}(A_s/q) \tag{2.5}$$

Para garantir uma qualidade mínima exige-se, na transmissão de voz digitalizada, uma relação sinal-ruído de quantificação para o PCM uniforme não inferior a 26 dB. Para além de proporcionar uma qualidade elevada para sinais de baixa amplitude um sistema telefónico deve ter capacidade para transmitir uma grande gama de amplitudes, o que se designa por gama dinâmica. A gama dinâmica é geralmente expressa como a relação em dB entre a amplitude máxima ($A_{s\,\max}$) e a sua amplitude mínima ($A_{s\,\min}$), ou seja

$$GD = 20\log_{10}\left(\frac{A_{s\max}}{A_{s\min}}\right) \tag{2.6}$$

Um valor típico para a gama dinâmica é 30 dB. Considere-se um quantificador uniforme, ou seja, um codificador em que a largura *q* de todos os intervalos de quantificação é a mesma. Se se designar por *L* o número de níveis de quantificação, então vem que

$$q = \frac{2A_{s\max}}{L} \tag{2.7}$$

onde $A_{s\max}$ é amplitude máxima do sinal que é possível quantificar sem distorção de corte. Inserindo (2.1) e (2.7) em (2.5) chega-se a

$$\frac{S}{N_q} = 1.76 + 6.02N_b + 20\log\left(\frac{A_s}{A_{s\max}}\right) \tag{2.8}$$

Como se conclui da Eq. (2.8), para um determinado número de bits por amostra constante, a relação sinal-ruído de quantificação aumenta com a amplitude A_s do sinal a quantificar. Isto significa que os sinais com baixa amplitude têm uma relação S/N_q baixa, enquanto os sinais com amplitude elevada apresentam S/N_q elevadas. Como os sinais com amplitudes elevadas têm pouca probabilidade de ocorrer conclui-se que o PCM uniforme é pouco eficiente.

Exemplo 2.2
Um sinal sinusoidal com 1 V de amplitude é digitalizado com uma relação sinal-ruído de quantificação mínima de 30 dB: Pede-se o número de intervalos de quantificação (quantificação uniforme) e o número de bits necessários para codificar cada amostra.

Solução: Usando a Eq.(2.5) conclui-se que q=0.078 V. Necessitam-se, assim, de 13 intervalos para codificar cada polaridade (1/0.078≈13), o que dá um total de 26. O número de bits é calculado usando (2.1) obtendo-se 5 bits por amostra.

Outro aspecto importante é que o ruído pode ser maior que o sinal quando os valores das amostras estão no primeiro intervalo de quantificação. Este efeito é particularmente perturbador durante as pausas de fala e é chamado **ruído do canal em repouso.**

A acção deste ruído nos sistemas PCM pode ser minimizada colocando um nível de quantificação em zero. Neste caso, todos os valores das amostras no intervalo de quantificação central são descodificados com uma saída constante a zero. Os sistemas deste tipo usam um número ímpar de intervalos de quantificação (ver Figura 2.6 b).

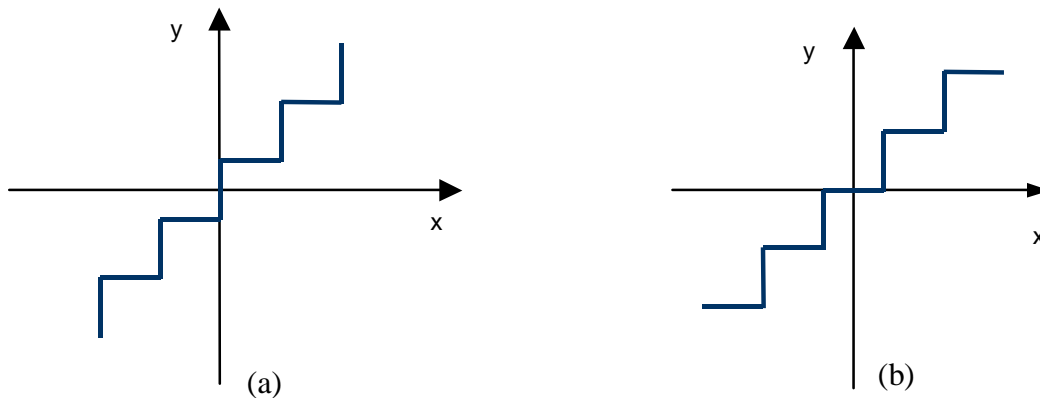



Figura 2.6 Característica de um quantificador. (a) Número par de níveis; (b) Número impar de níveis.

Os quantificadores com um número par de níveis (Figura 2.6a) não podem originar um nível de saída zero, enquanto o segundo descodifica os sinais fracos com uma saída de nível zero, como já se viu. Contudo, neste último caso, se a amplitude do sinal é semelhante à largura do intervalo de quantificação, ou se houver uma componente contínua no sinal, teremos também problemas de ruído do canal de repouso.

Um problema adicional ocorre quando a amplitude do sinal analógico ultrapassa os valores extremos do quantificador. Os sinais nessa situação dizem-se na **região de corte** e estão sujeitos a distorção de sobrecarga, a qual irá ser analisada na secção seguinte. Em conclusão, o comportamento de um quantificador pode-se equacionar identificando as seguintes regiões:

- Região de quantificação;
- Região de canal em repouso;
- Região de corte.

2.2.4.5 Distorção de sobrecarga

Na análise realizada nas secções anteriores considerou-se que o sinal de entrada tinha uma amplitude limitada ao intervalo $[-V,V]$, (onde $V=Lq/2$) também designada por gama de entrada autorizada. Qualquer sinal à entrada do quantificador que esteja fora desse intervalo é quantificado com o nível máximo permitido, ou seja, $\pm(V-q/2)$. Os sinais com essas características dizem-se que estão na **região de corte.**

Na presença de sinais com amplitudes superiores aos valores extremos o quantificador vai originar uma distorção de sobrecarga. Para calcular o valor médio quadrático dessa distorção (n_d) é necessário especificar a função densidade de probabilidade do sinal de entrada $p(x)$. Assumindo que a função densidade é simétrica então

$$n_d = 2\int_V^{\infty}(V-x)^2 p(x)dx \qquad (2.9)$$

Os sinais de voz são muitas vezes modelados por uma estatística Laplaciana, sendo neste caso a função densidade de probabilidade dada por:

$$p(x) = \frac{1}{\sigma_x\sqrt{2}}\exp(-\sqrt{2}|x|/\sigma_x) \qquad (2.10)$$

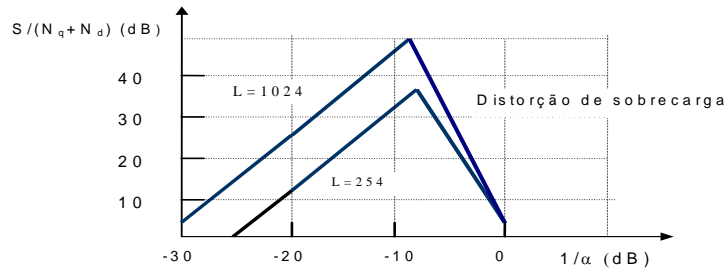
onde $\sigma_x^2 = s$ é a potência média do sinal. Entrando com (2.10) em (2.9) conclui-se que

$$n_d = \sigma_x^2 \exp\left(-\sqrt{2}V/\sigma_x\right) \qquad (2.11)$$

Exemplo 2.3 Admita que um sinal analógico com uma potência σ_x^2 é descrito por uma estatística Laplaciana é amostrado e, em seguida, quantificado por um quantificador uniforme com L níveis e uma gama autorizada de [-V,V]. Determine e represente a relação sinal-ruído de quantificação entrando em consideração com a distorção de sobrecarga.

Solução: Com base em (2.5), (2.7) e (2.11) chega-se a

$$\frac{s}{n_q + n_d} = \frac{\sigma_x^2}{\frac{V^2}{3L^2} + \sigma_x^2 \exp(-\sqrt{2}\frac{V}{\sigma})} = \frac{1}{\frac{\alpha^2}{3L^2} + \exp(-\sqrt{2}\alpha)}$$

onde $\alpha = V/\sigma_x$ é o factor de carga.

2.2.4.6 Quantificação não uniforme

Como já se referiu, os sinais de voz são caracterizados pelas pequenas amplitudes serem muito mais prováveis que as amplitudes mais elevadas. Deste modo, em vez de usar uma quantificação uniforme será mais razoável usar um quantificador com intervalos de quantificação largos para os sinais de amplitudes elevadas, e intervalos mais estreitos para amplitudes baixas, ou seja, **quantificação não uniforme.**

Com a quantificação não uniforme as pequenas amplitudes sofrem uma distorção inferior à das amplitudes mais elevadas, sendo possível garantir uma relação sinal-ruído de quantificação constante, desde que a característica de quantificação seja apropriada.

Para se poder determinar essa característica será necessário começar por se calcular a relação sinal-ruído de quantificação. Para isso, considere-se um sinal com amplitude normalizada *x*, situado no intervalo

$$x_i - 1/2\delta x_i \le x < x_i + 1/2\delta x_i \tag{2.12}$$

o qual é representado pela amplitude quantificada x_i. Seja $p(x)$ a função densidade de probabilidade da amplitude do sinal *x*. Se o incremento δx_i for pequeno em relação à gama do sinal, $p(x)$ varia pouco dentro do intervalo $[x_i - 1/2\delta x_i, x_i + 1/2\delta x_i]$ e pode ser aproximado por $p(x_i)$. Deste modo, o valor quadrático médio do erro de quantificação para os sinais situados no intervalo i é dado por:

$$< \varepsilon_i^2 >= \int_{x_i - 1/2\delta x_i}^{x_i + 1/2\delta x_i} (x - x_i)^2 p(x) dx = p(x_i) \int_{-1/2\delta x_i}^{1/2\delta x_i} \varepsilon_i^2 d\varepsilon_i = p(x_i) \frac{\delta x_i^3}{12} \tag{2.13}$$

Por sua vez a probabilidade que o sinal x esteja no i-ésimo intervalo é dada por

$$P_i = \int_{x_i - \delta x_i/2}^{x_i + \delta x_i/2} p(x) dx = p(x_i) \delta x_i \tag{2.14}$$

Substituindo este valor em (2.13) chega-se a

$$< \varepsilon_i^2 >= \delta x_i^2 P_i / 12 \tag{2.15}$$

O ruído de quantificação resulta das contribuições dos vários incrementos, ou seja

$$n_q = \sum_i < \varepsilon_i^2 >= \frac{1}{12} \sum_i \delta x_i^2 P_i \tag{2.16}$$

No caso específico da quantificação uniforme tem-se $\delta x_i = q$ o que substituído em (2.16) conduz à equação 2.3.

Seja y uma variável que descreve o sinal na saída do quantificador, que se admite normalizada relativamente ao valor da tensão máxima autorizada, cobrindo a gama ±1, enquanto x, que descreve o sinal normalizado na entrada, cobre a mesma gama. Deste modo, os intervalos uniformes sobre a escala dos y têm um valor de $2/L$, em que L é o número de níveis. Como a característica de quantificação não uniforme é descrita pela função $y=f(x)$ não é difícil demonstrar que

$$\left(\frac{dy}{dx} \right)_{x=x_i} = \frac{2}{L \delta x_i} \tag{2.17}$$

Usando (2.16) e (2.17) obtém-se a seguinte expressão para o ruído de quantificação não uniforme:

$$n_q = \frac{1}{3L^2} \sum_i P_i \left(\frac{dx}{dy} \right)_{x=x_i}^2 \tag{2.18}$$

a qual nos casos em que o número de níveis é suficientemente elevado, pode ser aproximada por

$$n_q = \frac{1}{3L^2} \int_{-1}^{1} p(x) \left(\frac{dx}{dy} \right)^2 dx \tag{2.19}$$

Como mostra a expressão anterior o ruído de quantificação depende da estatística do sinal analógico que se pretende codificar. A potência desse sinal também é dependente dessa estatística, já que

$$s = \langle x^2 \rangle = \int_{-1}^{1} x^2 p(x) dx \tag{2.20}$$

Com base nas equações anteriores conclui-se que a relação sinal-ruído de quantificação é dada por

$$\frac{s}{n_q} = 3L^2 \frac{\int_{-1}^{+1} x^2 p(x) dx}{\int_{-1}^{+1} \left(\frac{dx}{dy}\right)^2 p(x) dx} \tag{2.21}$$

Terá interesse em definir-se uma característica de quantificação tal que a relação sinal-ruído de quantificação seja independente da estatística do sinal de entrada. Isto consegue-se com

$$\frac{dx}{dy} = kx \tag{2.22}$$

onde k é uma constante. A relação sinal-ruído resultante é dada por

$$\frac{s}{n_q} = \frac{3L^2}{k^2} \tag{2.23}$$

A solução da equação diferencial (2.22) conduz-nos a uma característica de **compressão logarítmica**. Escolhendo a constante de integração tal que y=1, para x=1, essa característica ou lei de compressão é dada por

$$y = 1 + \frac{1}{k} \ln x \tag{2.24}$$

Como facilmente se conclui esta característica é irrealizável devido à sua assímptota vertical para x=0. Na secção seguinte analisar-se-ão técnicas que permitem ultrapassar este problema.

2.3 Codificação de fonte

A codificação de fonte designa vários tipos de processamento do sinal que permitem reduzir o débito binário dos sinais digitais, sem perda significativa de qualidade relativamente ao sinal original. Testes subjectivos demonstraram que para garantir uma boa qualidade de transmissão de voz é necessário garantir

$$\left(S/N_q\right)_{dB} \geq 35\,\text{dB} \tag{2.25}$$

Devido à dispersão estatística do nível médio da voz é necessário que essa condição seja respeitada não só para os níveis máximos do sinal, mas também num domínio da ordem de 40 dB abaixo desse nível. Com codificação uniforme tal exigira a utilização no mínimo de 12 bit/amostra, o que conduziria a um débito para um canal de telefonia de 96 kbit/s. Para garantir essa exigência com um débito de 64 kb/s é necessário a utilização de técnicas técnicas de quantificação não uniformes. Se se explorar a correlação existente entre as amostras de sinais, tais como os sinais de voz e de vídeo, é possível ainda garantir essas exigências com débitos inferiores, usando as denominadas técnicas de codificação diferencias

2.3.1 PCM não linear (leis A e μ)

Já vimos nas secções anteriores que associado ao processo de quantificação existe sempre ruído de quantificação, o qual, como processo irreversível que é, não pode ser eliminado por meios técnicos.

No PCM uniforme ou linear o efeito deste ruído pode tornar-se tão pequeno quanto se deseje, à custa do aumento do número de níveis de quantificação e do consequente débito binário.

Consegue-se, no entanto, reduzir o débito binário e manter uma relação sinal-ruído de quantificação desejável usando quantificação não linear. Essa quantificação é conseguida nos sistemas telefónicos usando uma característica de compressão logarítmica, como já se referiu, a qual é implementada nos sistemas práticos usando ou a lei μ (Japão, USA) ou a lei A (Europa). Considerando, novamente, amplitudes normalizadas a lei μ é definida por

$$y = \text{sgn}(x)\frac{\ln(1+\mu|x|)}{\ln(1+\mu)} \tag{2.26}$$

onde x é a amplitude do sinal de entrada $(-1 \leq x \leq 1)$, $\text{sgn}(x)$ é o sinal de x e μ é uma parâmetro usado para definir o grau de compressão, considerando-se usualmente μ=255. Por sua vez, a lei A é descrita por

$$y = \begin{cases} \text{sgn}(x)\left[\dfrac{A|x|}{1+\ln(A)}\right] & 0 \leq |x| \leq \dfrac{1}{A} \\ \\ \text{sgn}(x)\left[\dfrac{1+\ln|Ax|}{1+\ln(A)}\right] & \dfrac{1}{A} \leq |x| \leq 1 \end{cases} \tag{2.27}$$

com A=87.6. Na curva correspondente à lei A podemos distinguir duas regiões:

- Região de baixas amplitudes ($x \leq 1/A$) com variação linear;
- Região com amplitudes entre $1/A$ e 1 com variação logarítmica.

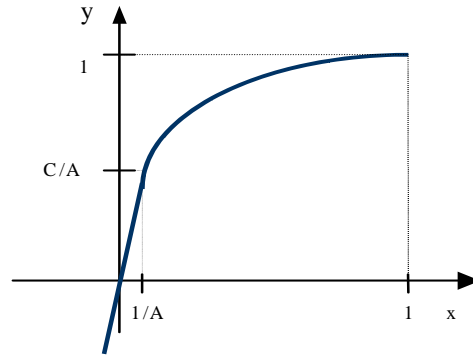



Figura 2.7 Característica de compressão segundo a lei A.

O declive do segmento linear na origem é definido por

$$C = \frac{A}{1 + \ln A} \tag{2.28}$$

o qual foi escolhido igual a 16, o que conduz ao valor de A acima referido (87.6). Usando as equações (2.21) e (2.27) pode-se obter a relação sinal-ruído de quantificação para a lei A (ver Figura 2.8), a qual se se considerar a parte linear fica

$$s/n_q = 3L^2C^2 < x(t)^2 > \tag{2.29}$$

Verifica-se que relativamente à quantificação uniforme com o mesmo número de L níveis a relação sinal-ruído de quantificação sofre uma melhoria de C^2, e por isso o parâmetro C^2 designa-se por **vantagem de compressão**. Obtém-se, desta maneira, para amplitudes reduzidas a mesma qualidade que se obteria com uma quantificação uniforme C=16 vezes mais fina, ou seja, com 16L níveis. Por sua vez a relação sinal-ruído para a parte logarítmica é dada por

$$s/n_q = 3L^2C^2 / A^2 \tag{2.30}$$

Aqui a relação sinal-ruído de quantificação é constante e independente da amplitude do sinal. Assim, para $< x(t)^2 >= 1$, a relação sinal-ruído reduz-se de um factor igual a A^2 / C^2 relativamente ao caso uniforme. Para garantir a condição (2.25) o ITU-T fixou um número de níveis de quantificação igual a 256 (o que implica a utilização de um código binário com palavras de 8 bits). Atendendo a que o ritmo de amostragem para os sinais telefónicos também foi fixado por convenção internacional igual a 8 kHz, chega-se a débito binário de 64 kb/s, como já tinha sido referido anteriormente. Para a lei μ (μ=255) o declive na origem é dada por

$$C = \frac{\mu}{\ln(1 + \mu)} \cong 46 \tag{2.31}$$

o que permite concluir que a lei μ conduz a uma ligeira melhoria da relação sinal-ruído relativamente à lei A, para sinais com amplitudes reduzidas.

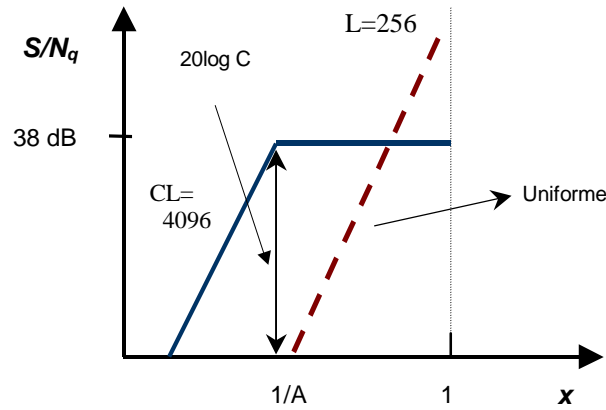



Figura 2.8 Relacão sinal-ruído de quantificação para a lei A.

A vantagem de compressão como já foi visto descreve a melhoria obtida com a quantificação não uniforme para sinais fracos. Um outro indicador de desempenho de índole mais geral é o **ganho de compressão**. Este parâmetro é definido como a razão entre a relação sinal-ruído da quantificação não uniforme e a relação sinal-ruído da quantificação uniforme. Atendendo a (2.21) esse ganho é dado por

$$g_c = \frac{1}{\int_{-1}^{+1} \left(\frac{dx}{dy} \right)^2 p(x) dx} \tag{2.32}$$

2.3.2 Técnicas para realizar a codificação não uniforme

A realização de uma quantificação não uniforme pode ser combinada com a codificação através de um dos três métodos apresentados na Figura 2.9. A Figura 2.9 (a) representa o método discutido até agora, o qual é baseado na compressão analógica seguida de uma quantificação uniforme com 8 bits. O principal problema deste método consiste em garantir a reciprocidade exacta das características de compressão e expansão. No método da Figura 2.9 (b) tem-se uma quantificação uniforme com *CL* níveis, o que corresponde a $\log_2 CL$=12 bits, seguida de uma compressão digital (tradução digital) reduzindo a dimensão das palavras a 8 bits. No caso da Figura 2.9 (c) tem-se uma codificação não uniforme com 8 bits com a característica de compressão incorporada no codificador. Neste caso, a lei de compressão é aproximada por uma característica segmentada e, deste modo, perfeitamente adaptada ao processamento digital.

A lei A usa 13 segmentos, com sete positivos e sete negativos, sendo os dois segmentos que passam pela origem colineares e, por isso, são contabilizados como um único segmento designado por central. O segmento central tem um declive igual a 16, o que corresponde ao declive dado por (2.28). Os declives dos segmentos estão em

progressão geométrica com razão igual a 1/2 (ver Figura 2.10). No interior de cada segmento a quantificação é uniforme com 16 níveis (32 no segmento central). Como resultado, quando o sinal aumenta, o passo de quantificação também aumenta de um factor 2 de um segmento para o seguinte.

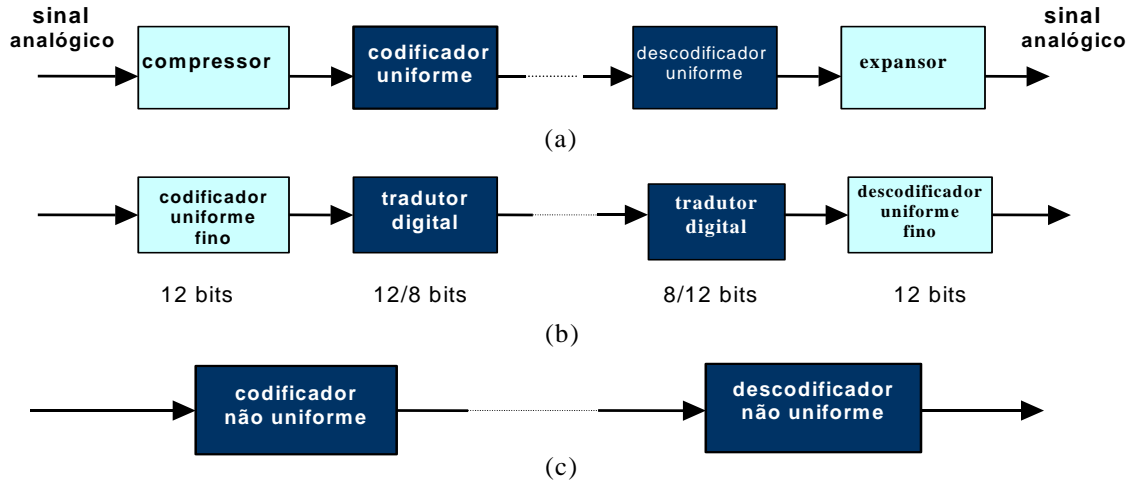



Figura 2.9 Técnicas de codificação não uniformes.

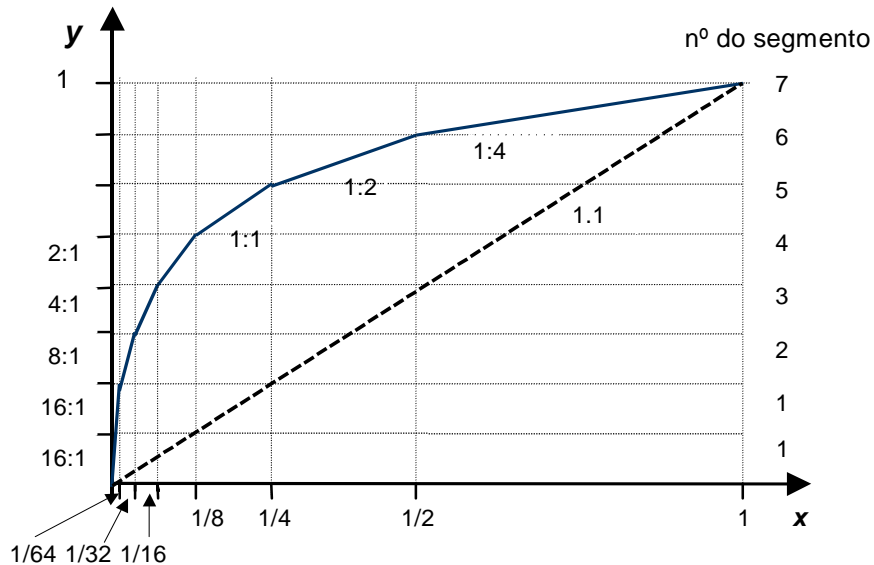



Figura 2.10 Característica de compressão digital com 13 segmentos (valores positivos).

Cada palavra PCM de 8 bits é codificada do seguinte modo: o primeiro bit indica a polaridade, os três bits seguintes identificam o segmento (000 e 001 correspondem ao segmento central) e os quatros últimos bits o nível de quantificação do conjunto de 16. A lei μ adopta uma solução idêntica, usando, contudo, 15 segmentos em vez dos 13 de lei A. Na tabela 2.1 apresentam-se os pontos extremos dos segmentos, os intervalos de quantificação e as correspondentes palavras de código.

Tabela 2.1 Tabela de codificação de lei A segmentada

número do segmento	gama do sinal	dimensão do passo	código do segmento	código de quantificação
1	0-2 2-4 . 30-32 32-34 . 62-64	2	000 001	0000 0001 . 1111 0000 . 1111
2	64-68 . 124-128	4	010	0000 . 1111
3	128-136 . 248-256	8	011	0000 . 1111
4	256-272 . 496-512	16	100	0000 . 1111
5	512-544 . 992-1024	32	101	0000 . 1111
6	1024-1088 . 1984-2048	64	110	0000 . 1111
7	2048-2176 . 3968-4096	128	111	0000 . 1111

2.3.3 Algoritmo de codificação para a lei A segmentada

A estrutura de uma palavra PCM (lei A), como já se referiu, usa um bit para identificar a polaridade da amostra (bit *P*) e os restantes sete para especificar a sua amplitude. Três dos bits de amplitude são usados para especificar o segmento e são caracterizados pelo identificador *S,* enquanto os outros quatro são usados para identificar o intervalo dentro de cada segmento, e são caracterizados pelo identificador *Q*. O bit de polaridade é igual a zero para valores positivos das amostras do sinal e igual a 1 para valores negativos (ver Figura 3.13).

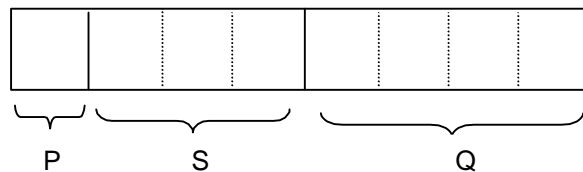



Figura 2.11 Estrutura de uma palavra PCM (lei A).

Dada uma amostra com amplitude x (amplitude máxima admissível pelo codificador é de 4096) a primeira etapa consiste em determinar o identificador do segmento S. Os diferentes segmentos são identificados pelos seus pontos extremos finais os quais na lei A são 32, 64, 128, 256, 512, 1024, 2048 e 4096. Assim, o identificador S pode ser determinado como sendo o menor valor de a que verifica a relação

$$x < 32 \times 2^{a} \quad a=0,1,\ldots,7 \tag{2.33}$$

Depois de o segmento que contém a amostra ter sido identificado é necessário obter o intervalo de quantificação dentro desse intervalo. Como primeiro passo, obtém-se o resíduo r, ou seja, a diferença entre a amplitude da amostra e o ponto extremo inicial do segmento:

$$r = \begin{cases} x & S = 0 \\ x - 16 \times 2^{S} & S = 1,2,\ldots\ldots,7 \end{cases} \tag{2.34}$$

O valor de Q corresponde agora ao intervalo de quantificação que contém o resíduo r, sendo igual ao menor valor de b que verifica a relação:

$$\begin{aligned} r &< 2(b+1) & S &= 0 \\ &< 2^{S}(b+1) & S &= 1,2,\ldots\ldots,7 \end{aligned} \tag{2.35}$$

Depois de S e Q serem determinados, estes são codificados, respectivamente, com três e quatro bits. Estes bits são concatenados e formam uma palavra de 7 bits, a qual corresponde um inteiro entre 0 e 127, que, em síntese, identifica um dos 128 intervalos de quantificação associados a cada polaridade.

O processo de descodificação compreende a atribuição da polaridade apropriada a uma amostra analógica, que tem uma amplitude igual a metade do intervalo de quantificação de ordem n (n=0, 1,….., 127). Com base nos valores de S e Q obtidos, pode-se determinar a amplitude dessa amostra usando a equação

$$y_n = \begin{cases} 2Q + 1 & S = 0 \\ 2^{S}(Q + 16.5) & S = 1,2,\ldots.,7 \end{cases} \tag{2.36}$$

Exemplo 2.4

Considere uma amostra x=-100, determine a palavra de código correspondente e o valor da saída do descodificador.

Solução: Usando o algoritmo de codificação referido, determina-se P=1, S=2 e Q=9, o que corresponde à palavra 1 010 1001. Essa palavra corresponde a uma amostra que se encontra no intervalo de quantificação 41 (polaridade negativa), de modo que a saída do descodificador é dada por

$$y_{41} = -4(9 + 16.5) = -102$$

que corresponde a metade do intervalo entre -104 e -100.

Em síntese, por acordo internacional (Recomendação G 711 do ITU-T) foram fixados os seguintes parâmetros para a representação digital dos sinais vocais telefónicos:

- frequência de amostragem de 8 kHz;
- quantificação não uniforme com 256 níveis;
- palavras PCM com 8 bits;
- débito binário para um canal de 64 kb/s;
- lei de codificação europeia: lei A com 13 segmentos;
- lei de codificação americana: lei μ com 15 segmentos.

2.3.4 PCM diferencial

Em certos tipos de sinais, como, por exemplo, a voz e o vídeo o valor do sinal varia pouco de amostra para amostra. Tal comportamento permite estimar (extrapolar) o valor de uma amostra a partir das amostras anteriores e transmitir unicamente a informação correspondente à diferença entre os valores estimados e os valores actuais. Como a gama de variação das diferenças é menor que a das amostras, são necessários menos bits para proceder à sua codificação e por que por conseguinte vai conduzir a uma redução do débito binário,

As técnicas de modulação diferenciais baseiam-se nessas propriedades sendo, por conseguinte, tanto mais eficazes quanto maior a redundância dos sinais a processar e melhor o conhecimento das suas leis estatísticas.

A modulação PCM diferencial ou DPCM (*differential pulse code modulation*) foi desenhada especificamente para aproveitar a redundância presente nas diferentes amostras do sinal de voz. Com o DPCM é transmitida a diferença entre as amplitudes de uma amostra e de uma estimativa da amostra anterior.

Para além das operações presentes nos codecs de PCM é necessário acrescentar no DPCM um operação adicional designada por predição (*prediction*), pela qual se obtém uma estimativa da amostra anterior. O preditor pode tomar a forma de um filtro transversal, do tipo representado na Figura 2.16. A função de transferência deste filtro é dada por

$$H(f) = \sum_{k=-\infty}^{\infty} h(k) e^{-j(2\pi fkT)} = \sum_{k=0}^{n} c_k e^{-j(2\pi fkT)} \tag{2.37}$$

onde $h(k)$ é a resposta impulsiva do filtro e c_k o coeficiente de ordem k do filtro.

Exemplo 2.5

Um filtro transversal é caracterizado pelos seguintes coeficientes:

$$h(k) = a^k u(k) \quad |a| < 1$$

onde $u(k)$ representa a função escalão. Pretende-se terminar a função de transferência desse filtro.

Solução: De acordo com (2.37) tem-se

$$H(f) = \sum_{k=-\infty}^{\infty} a^k e^{-j(2\pi fkT)} u(k) = \sum_{k=0}^{\infty} \left[a e^{-j(2\pi fT)} \right]^k = \frac{1}{1 - a e^{-j(2\pi fT)}}$$

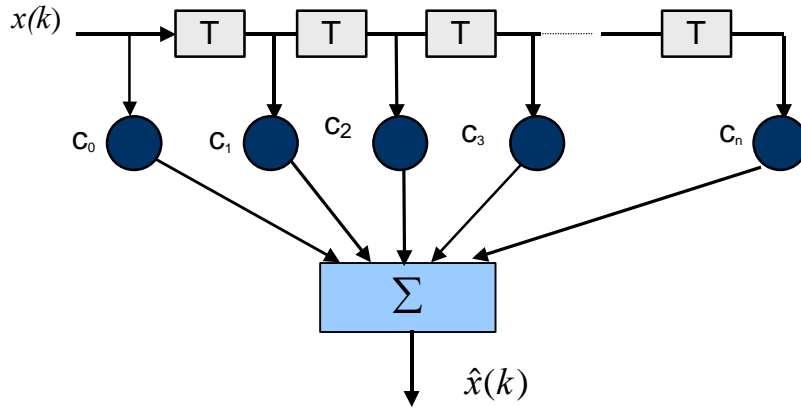



Figura 2.12 Estrutura de um filtro transversal usado como preditor.

Atendendo à definição de convolução discreta a saída desse filtro $\hat{x}(k)$, relaciona-se com a sua entrada $x(k)$, através de

$$\hat{x}(k) = \sum_{i=-\infty}^{\infty} h(i)x(k-i) = \sum_{i=0}^{n} c_i x(k-i) \tag{2.38}$$

A ordem de predição tem a ver com o número de coeficientes do filtro. Por exemplo, um preditor de terceira ordem é constituído por um filtro transversal com três coeficientes.

O esquema de blocos de um modulador DPCM está representado na Figura 2.13. De acordo com essa figura o sinal diferença $x(k)-\hat{x}_q(k)$ é aplicado à entrada do quantificador, sendo a sua saída codificada tal como no PCM binário.

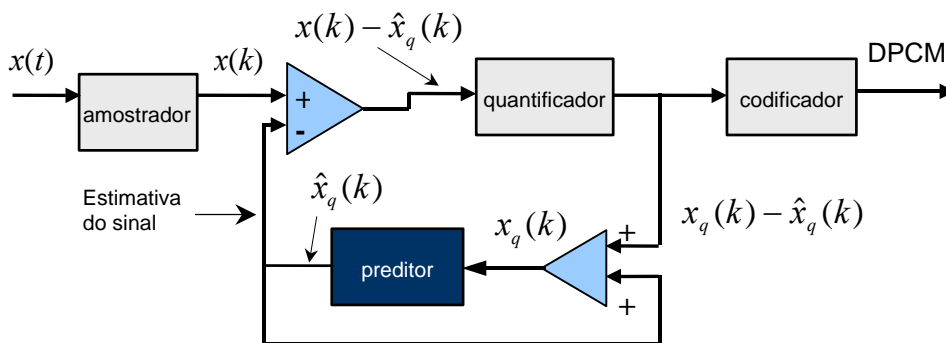



Figura 2.13 Esquema de blocos de um modulador DPCM.

Um das dificuldades deste esquema prende-se com a codificação de sinais com variações de amplitude muito rápidas. Para garantir que nestas situações a distorção introduzida no sinal não é significativa deve garantir-se que declive do quantificador deve ser superior ou igual à variação máxima do sinal de entrada, o que se traduz por

$$\left|\frac{dx}{dt}\right|_{\max} \leq f_s(L-1)\Delta \qquad (2.39)$$

onde f_s designa a frequência de amostragem, L o número intervalos de quantificação e 2Δ o espaçamento entre esses intervalos.

Um parâmetro usado, normalmente, para quantificar o desempenho de um codificador DPCM é o ganho de predição, g_p. Este ganho é definido como sendo o valor de melhoria da relação sinal-ruído de quantificação do DPCM relativamente ao PCM, ou seja

$$\left(\frac{s}{n_q}\right)_{DPCM} = g_p\left(\frac{s}{n_q}\right)_{PCM} \qquad (2.40)$$

com

$$g_p = \left[1-\sum_{i=1}^{n} c_i\rho_i\right]^{-1} \qquad (2.41)$$

onde $\rho_i = R_x(iT)/\sigma_x^2$ é o coeficiente de correlação do sinal de entrada. Os coeficientes do filtro transversal são calculados com base nesses coeficientes, do seguinte modo:

$$\begin{bmatrix} \rho_0 & \rho_1 & .. & \rho_{n-1} \\ \rho_1 & \rho_o & .. & \rho_{n-2} \\ . & . & & . \\ \rho_{n-1} & \rho_{n-2} & .. & \rho_o \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ . \\ c_n \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ . \\ \rho_n \end{bmatrix} \qquad (2.42)$$

No caso da voz, os coeficientes do filtro são calculados usando as características médias do sinal de voz. Isto conduz a ganhos de predição entre 5 a 10 dB. Para o caso do vídeo, como a correlação entre as amostras é mais elevada, podem-se obter ganhos da ordem de 12 dB.

Uma das desvantagens do DPCM reside no facto de se as características do sinal de entrada se afastarem das características médias referidas poder ter-se uma degradação significativa do desempenho. Por exemplo, os sons vozeados têm uma característica de correlação elevada, enquanto os não vozeados é próxima de zero, o que faz com que a transmissão DPCM para estes últimos seja muito degradada.

Uma solução para este problema consiste em ajustar permanentemente os coeficientes do extrapolador às características do sinal de entrada. Somos, assim, conduzidos ao **DPCM adaptativo ou ADPCM**, que se baseia não só na adaptação do preditro, mas também no ajuste do número de níveis do quantificador. Com esta técnica é possível, por exemplo, a transmissão de voz a 32 kb/s, com uma qualidade praticamente idêntica à do PCM a 64 kb/s e se for aceitável alguma redução na qualidade pode mesmo chegar-se aos 16 kbit/s.

2.4 Outros tipos de codificação

A maior parte das técnicas de codificação descritas têm como objectivo principal a reprodução tão fiel quanto possível da forma de onda do sinal transmitido. Em geral, como não estão condicionadas pela natureza do sinal que processam, são aplicáveis a qualquer tipo de sinal presente no canal de voz.

Existem, contudo, uma outra família de codificadores que são específicos para sinais de voz, designando-se por codificadores de voz ou *vocoders* (*voice coders*). Este tipo de codificadores não pode, no entanto, ser aplicado na rede telefónica pública, onde para além dos sinais de voz se têm outros sinais, como por exemplo, os sinais gerados pelos modems.

A ideia que subjaz aos *vocoders* é a transmissão de um certo número de parâmetros característicos do sinal de voz que possibilitem a sua futura síntese no receptor. Podem-se, assim, identificar duas fases no processo de codificação: análise ou geração desses parâmetros no emissor e síntese do sinal no receptor.

Existem vários tipos de *vocoders.* Aqui, iremos, simplesmente, tecer algumas considerações sobre a técnica LPC (*Linear Predictive Coding*), que é usada não só nos *vocoders*, mas também nos codificadores híbridos. No âmbito dos *vocoders*, esta técnica de codificação é uma boa solução, sempre que se requerem débitos entre 1.2–2.4 kbit/s e a qualidade da voz não é um factor determinante.

2.4.1 LPC

O processo de análise do sinal de voz realizada por um codificador LPC permite modelar a função de transferência do tracto vocal e o tipo de excitação associados aos sinais de voz. A função de transferência é descrita por um filtro transversal, o que, como já se viu, implica a definição de um conjunto de coeficientes. Para além destes, é necessário definir um termo de ganho, associado à amplitude do sinal de voz. A caracterização do tipo de excitação envolve a definição do tipo dos sons (vozeados ou não vozeados) e no caso do vozeados do seu pseudo-período ou *pitch.*

Esses parâmetros, que são actualizados periodicamente, são codificados e transmitidos. Com base neles, no terminal receptor um sintetizador reproduz o sinal de voz original. A estrutura de um sintetizador é apresentada na Figura 2.14. Os sinais vozeados são reproduzidos ligando um gerador de pulsos com o pseudo-período referido, enquanto os não vozeados requerem a utilização de um gerador de ruído branco. No caso específico do *vocoder* DOD LPC 10 o filtro transversal é constituído por 10 coeficientes, a codificação é feita usando uma trama constituída por 180 amostras e com duração de 22.5 ms[23]. O número de bits requeridos por trama é igual a 54, que são repartidos do seguinte modo: 42 bit para a informação dos coeficientes do filtro; 7 bit para o pseudo-período e para a informação de voz; 5 bit para o ganho do amplificador. Nestas condições é fácil verificar que o débito binário deverá ser igual a 2.4 kbit/s, tendo-se assim, conseguido compressão do débito de cerca de 27 vezes relativamente aos valores obtidos com o PCM convencional.

---

[23] O período de amostragem é de 125 μs (inverso da frequência de amostragem), o que conduz a 125 μs×180= 22.5 ms

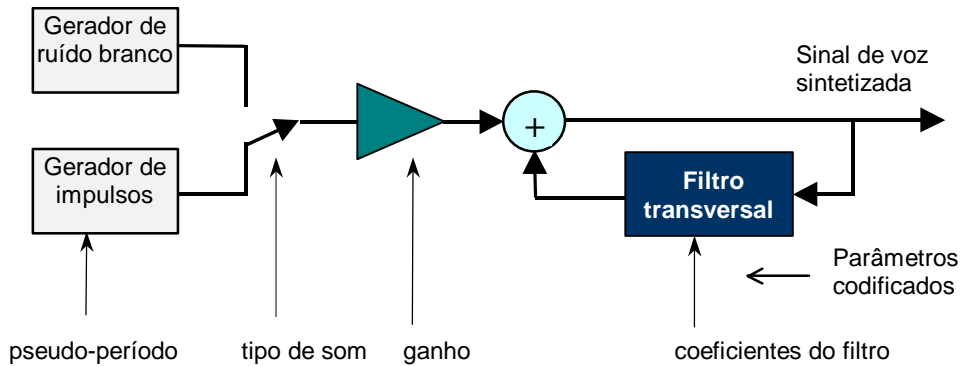



Figura 2.14 Sintetizador de voz usado na codificação LPC.

2.4.2 Qualidade dos codificadores

Nas secções anteriores estudaram-se, ainda que de modo algo sumário dois tipos de codificadores apropriados para comprimir o débito dos sinais vocais:

- Os codificadores de forma de onda, que se baseiam no facto das amostras sucessivas estarem altamente correlacionadas e que usam técnicas de codificação diferencial e quantificação adaptativa, as quais permitam obter bons desempenhos para débitos entre 24 kbit/s e 32 kbit/s.
- Os codificadores LPC que usam técnicas baseadas na síntese de voz, permitem débitos entre 1.2 kbit/s e 2.4 kbit/s, mas a qualidade da voz é muito baixa.

Os **codificadores híbridos** usam componentes das duas famílias especificadas anteriormente e permitem obter esquemas de codificação muito eficientes para operação entre 6 kbit/s e 16 kbit/s.

A qualidade dos codificadores usados para comprimir o débito binário dos sinais de voz é, normalmente, avaliada usando um método de medida subjectivo, designado por MOS (*mean opinion score*). Neste tipo de método é pedido aos intervenientes num teste para classificarem a qualidade de diferentes amostras de sinais de voz, sem terem conhecimento da origem dessas amostras. A qualidade dessas amostras é, geralmente, classificada usando a escala mostrada na Tabela 2.2. Os resultados MOS são calculados a partir das classificações atribuídas por um número suficientemente elevado de intervenientes de modo a terem significado estatístico.

Tabela 2.2 Classificações MOS

Qualidade	Classificação
Excelente	5
Bom	4
Razoável	3
Medíocre	2
Má	1

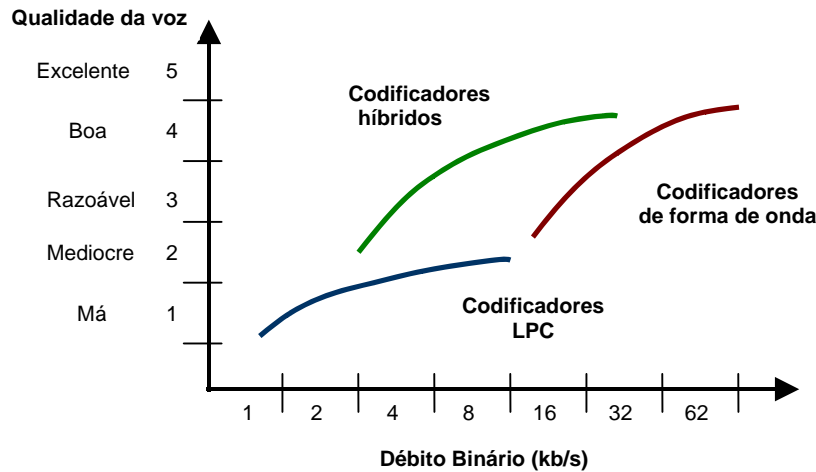



Figura 2.15 Classificações MOS de codificadores em função do débito binário.

A Figura 2.15 dá uma ideia das classificações MOS de diferentes tipos de codificadores e a tabela 2.3 apresenta mais detalhe sobre as características desses codificadores. O codec G.711 que corresponde ao PCM a 64 kbit/s, juntamento com o G.726 que usa ADPCM a 40 kbit/s são aqueles que proporcionam uma melhor qualidade. No ADPCM quando o débito decresce para 32 kbit/s a qualidade degrade-se ligeiramente, conduzindo de modo algo surpreendente a um desempenho idêntico aquele que é proporcionado pelo codec G.729, que opera com um débito de 8 kbit/s. Este último codec apresenta-se, assim, como uma solução bastante interessante, para aplicações em que um débito baixo e uma qualidade boa estão a prémio, como é o caso das aplicações de voz sobre IP. Em último lugar da tabela aparece o LPC10, cujos parâmetros foram apresentados anteriormente, e cuja qualidade é considerada má.

Tabela 2.3 Caracterização de diferentes tipos de codificadores.

Norma	G.711 (ITU-T)	G.726 (ITU-T)	G.723.1 (ITU-T)	G.729 (ITU-T)	DOD FS1015 (USA)
Tipo de codec	Forma	Forma	Híbrido	Híbrido	Vocoder
Débito (kbit/s)	64	16/24/32/40	5.3	8	2.4
Codificação	PCM	ADPCM	ACELP[24]	CS-ACELP[25]	LPC10
Qualidade MOS	4.2	2/3.2/4/4.2	3.7	4	Sintética

2.5 Voz sobre IP

No serviço de voz sobre IP (VoIP) o sinal de voz é segmentado em pacotes que são transportados como datagramas IP. O IP é um protocolo de nível 3 e contem soluções de endereçamento e de controlo apropriadas para encaminhar os pacotes entre a fonte e o

[24] *Algebraic code excited linear prediction*
[25] *Conjugate structure* ACELP

destino. Uma rede IP proporciona um serviço de entrega de datagramas não orientado à ligação e baseado no melhor esforço. Na sua forma mais simples uma rede VoIP é constituída por dois ou mais portais VoIP (*gateways*) interligados por uma rede IP. Estes dispositivos são responsáveis pela conversão dos sinais de voz num fluxo de pacotes IP, pela expedição desses pacotes para o destino e na recepção pela conversão do fluxo de dados IP de novo em sinais de voz.

O TCP (*Transmission Control Protocol*) e o UDP (*User Datagram Protocol*) são os protocolos de nível 4 usados, normalmente, numa rede IP. O primeiro é um protocolo orientado à ligação e oferece um serviço fiável e sem perdas. Um serviço sem perdas significa que ou todos os dados são entregues no destino, ou em caso de perdas o expedidor é informado da situação. O segundo, por sua vez, é um protocolo não orientado à ligação e baseado no melhor esforço, não oferecendo fiabilidade nem um serviço sem perdas. O protocolo UDP tem, contudo, a vantagem de requerer um cabeçalho mais reduzido do que o TCP (8 versus 20 octetos) e garante que o processo de expedição de dados é mais rápido. Devido a estas vantagens este protocolo é mais apropriado para aplicações em tempo-real como é o caso da voz sobre IP.

Das diferentes normas que estão disponíveis para o transporte de voz sobre IP/UDP a norma mais usada é aquela que corresponde à recomendação do ITU-T H.323. Esta norma, inclui não só especificações para a transmissão da informação do utilizador, mas também para a transmissão na rede IP de sinais de controlo e de sinalização adequados para iniciar, manter e terminar as ligações telefónicas (ver Figura 2.16). Para compensar o facto do protocolo UDP não garantir um serviço fiável e sem perdas, a recomendação H.323 propõe a utilização adicional do protocolo RTP (*Real-Time Protocol*) do modo como se descreve na Figura 2.16. O objectivo é garantir com esse protocolo um serviço que permita associar a cada pacote o tipo de utilização (voz, vídeo, etc.), que permita a numeração de sequência de modo a ser possível ordenar na recepção os pacotes que chegam fora de ordem ou atrasados, que permita a monitorização das entregas e que possibilite ainda a transmissão do sinal de relógio (*timestamping*). Esta última funcionalidade é deveras importante, pois os codecs do receptor para funcionaram têm de ter a informação de sincronismo enviada pelo codec emissor.

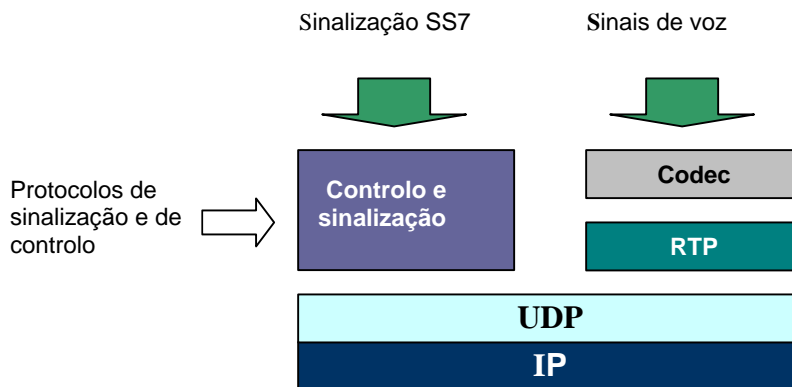



Figura 2.16 Protocolos usados no serviço de voz sobre IP.

Na figura 2.17 é detalhado o processo de formação dos pacotes no serviço de voz sobre IP. A maioria dos codificadores é orientada para a trama: isto quer dizer que a compressão é feita sobre conjuntos de amostras (tramas) em vez de ser feita trama a trama. É claro que o processo de formação de tramas requer um certo tempo, que vai desde os 10 ms para o codec G729 até 30 ms para o codec G.723.1, o qual se vai adicionar ao tempo de atraso extremo-a-extremo da ligação. Cada pacote pode transportar uma trama ou várias tramas. A primeira solução reduz o tempo de atraso mas também reduz a eficiência de transmissão, porque aumenta o número de cabeçalhos. A segunda solução aumenta essa eficiência, mas em contrapartida também aumenta o tempo de atraso. O impacto dos cabeçalhos associados aos diferentes protocolos é significativo. Por cada pacote têm-se 40 octetos de cabeçalho distribuídos do seguinte modo: 20 octetos para o IP[26], 8 octetos para o UDP e 12 octetos para o RTP. Tal implica, por exemplo, que no caso do codec G729, o débito na saída do codificador é de 8 kbit/s e depois da introdução dos cabeçalhos este débito aumenta para 24 kbit/s.

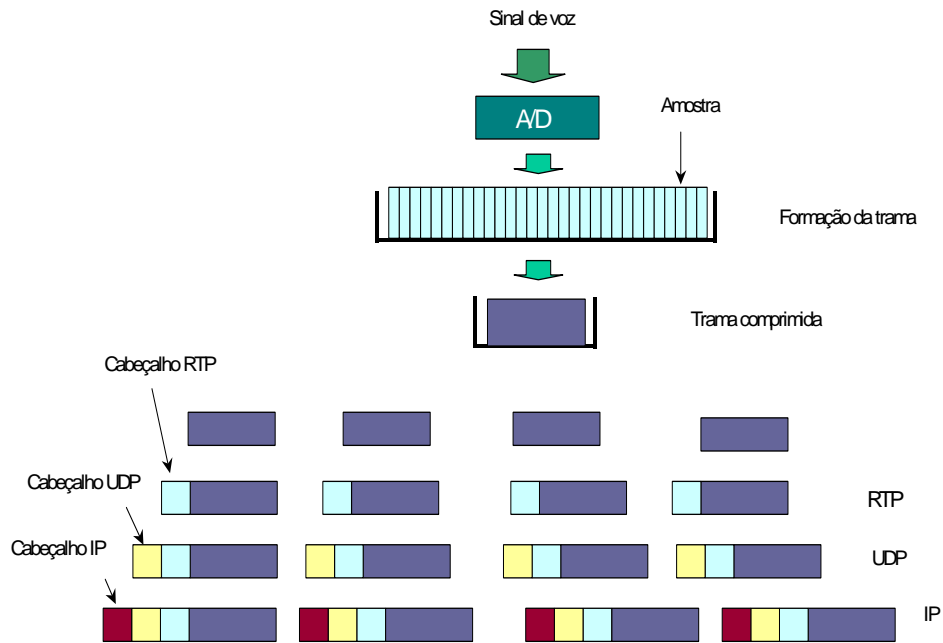



Figura 2.17 Processo de formação dos pacotes IP no serviço VoIP.

2.6 Problemas

2.1 A qualidade mínima requerida para a voz digitalizada exige uma relação S/N_q de pelo menos 26 dB. Admitindo que a gama dinâmica do sinal de voz é de 30 dB, calcule o número mínimo de bits necessários para garantir essa qualidade, usando quantificação uniforme.

[26] No IPv6 o número de octetos por cabeçalho aumenta para 40.

2.2 Qual é o débito binário de um sinal de alta-fidelidade digitalizado (PCM uniforme) com uma gama dinâmica de 40 dB? Suponha que os requisitos de fidelidade exigem a passagem de uma frequência de 10 kHz, com uma relação sinal-ruído de quantificação de 50 dB. Para simplificar suponha que os sinais de entrada são sinusoidais.

2.3 Um sistema de aúdio digital CD (*compact-disc*) usa PCM linear com 16 bits e uma frequência de amostragem de 44.1 kHz, para cada um dos dois canais estéreo. a) Qual é o débito binário resultante? b) Qual é a máxima frequência permitida no sinal de entrada? c) Qual é o valor máximo da relação sinal-ruído de quantificação em dB? d) Se os sinais musicais apresentarem um factor de carga de 20, qual é o valor médio da relação sinal-ruído de quantificação? e) Se o tempo de leitura do CD for de 70 minutos, determine o número total de bits armazenados no disco. Assuma que o código corrector de erros e a sincronização ocupam um quarto da capacidade total.

2.4 Um sinal de televisão tem uma largura de banda de 5.5 MHz. Qual é o débito binário requerido admitindo que este sinal é digitalizado com PCM uniforme e uma relação sinal-ruído de quantificação de 35 dB? Considere que a relação ritmo de amostragem-ritmo de Nyquist é a mesma que a usada na codificação da voz.

2.5 A largura de banda de um sinal de televisão é de 4.5 MHz. Este sinal é convertido para PCM linear com 1024 intervalos de quantificação. A frequência de amostragem está 20% acima do ritmo de Nyquist. a) Determine o débito binário resultante. b) Determine a relação sinal-ruído de quantificação se o factor de carga do quantificador é igual a 6.

2.6 Um codificador de PCM de 8 bits usando um quantificador de lei μ apresenta uma vantagem de compressão de 15 dB. Calcule a relação sinal/ruído de quantificação admitindo que a tensão do sinal está uniformemente distribuída.

2.7 Considere as seguintes palavras de código da lei A: 1 110 0010 e 0 010 1000. Determine a palavra de código correspondente à soma dessas duas palavras.

2.8 Considere um quantificador de lei A. a) Calcule os limites superior e inferior (independentes da estatística do sinal) da relação sinal/ruído de quantificação. b) Calcule a vantagem de compressão.

2.9 Considere uma lei de quantificação PCM segmentada com μ =255 (15 segmentos). Determinar a palavra de código que representa um sinal de 5 V se o codificador for projectado para uma gama de entrada de $\pm$ 10 V. Qual o valor da tensão depois da descodificação?

2.10 Considere um sistema DPCM com um filtro extrapolador de 2ª ordem. Determine o valor dos coeficientes do filtro, o ganho de extrapolação e o número de bits

economizados relativamente ao PCM binário, para os seguintes sinais: a)Voz com ρ_1=0.85 e ρ_2=0.64; b) Vídeo com ρ_1=0.95 e ρ_2=9.

2.7 Referências

[Be91] J. Bellamy, *Digital Telephony*, John Wiley & Sons, 1991.

[Ca86] A. B. Carlson, *Communication Systems*, McGraw-Hill, 1986.

[He05] O. Hersent, J. P. Petit, D. Gurle, *Beyond VOIP Protocols,* John Wiley&Sons, Ltd., 2005.

[Me01] P. Mehta, S. Udani,"Voice over IP", *IEEE Potencials*, pág. 36-40, Out./Nov. 2001.

[Si05] D. R. Smith, *Digital Transmission Systems*, 3ª edição. Kluwer Academic Publishers, 2004.

[Ta86] H.Taub, D. L. Schiling, *Principles of Communication Systems*, McGraw-Hill, 1986.

3 Aspectos de Transmissão

3.1 Suportes de transmissão

O meio de transmissão dominante na rede de acesso das redes telefónicas públicas é o cabo de pares simétricos, que como já se referiu é baseado em pares de cobre entrelaçados (daí essa designar-se rede de cobre). Por sua vez, na rede de transporte usa-se ou a fibra óptica ou meios radio-eléctricos, como os feixes hertzianos e os satélites. O cabo coaxial teve o seu período áureo na transmissão analógica, antes do aparecimento da tecnologia óptica. Contudo, a crescente implantação das redes de distribuição de televisão por cabo, baseadas na tecnologia híbrida (fibra/coaxial), fizeram ressurgir a importância deste meio de transmissão.

Nesta secção ir-se-ão analisar as principais características de transmissão (atenuação, distorção e diafonia) dos cabos de pares simétricos, dos cabos coaxiais e das fibras ópticas. Os meios de transmissão radio-eléctricos serão estudados em outras disciplinas.

3.1.1 Cabos de pares simétricos

Um par simétrico não é mais do que uma linha de transmissão constituída por dois condutores isolados e entrelaçados. O material usado nos condutores é, normalmente, o cobre, enquanto como isolador se usa o polietileno. No sentido de melhorar as propriedades de diafonia os quatro fios correspondentes a dois pares são entrelaçados formando uma **quadra,** designada por quadra-estrela. Um outro tipo, designada por quadra DM (do nome dos seus inventores Dieselhorst-Martin), ou quadra de pares combinados, é obtida entrelaçando dois pares previamente entrelaçados. Um cabo de pares simétricos é constituído por várias quadras dispostas de modo conveniente.

3.1.1.1 Parâmetros característicos de uma linha

A análise das características de transmissão dos pares simétricos pode-se fazer recorrendo à teoria das linhas de transmissão. De acordo com esta teoria, um troço elementar de linha pode-se modelar pelo esquema equivalente representado na Figura 3.1. Os elementos desse esquema designam-se por parâmetros primários da linha e são os seguintes:

- R (Ω/km): resistência por unidade de comprimento dos condutores da linha, incluindo o efeito pelicular;
- L (H/km): indutância por unidade de comprimento devida ao campo magnético entre os condutores;
- G(S/km): condutância transversal por unidade de comprimento devida, quer aos defeitos de isolamento, quer às perdas dieléctricas;
- C (F/km): capacidade por unidade de comprimento ligada à permitividade do dieléctrico situado entre os condutores.

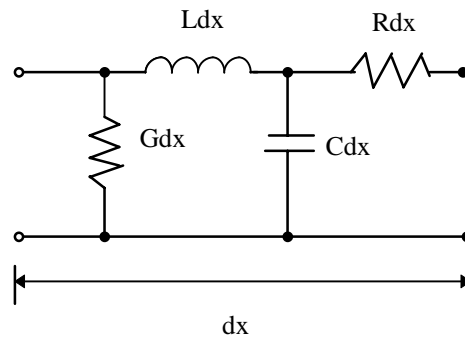



Figura 3. 1 Modelo de um troço elementar de linha de transmissão.

Tabela 3.1 Parâmetros primários característicos de um cabo do tipo 24 AWG(70 ºF)

f(kHz)	R(Ω/km)	L(mH/km)	G(μS/km)	C(μF/km)
1	172	0.613	0.071	0.052
5	172	0.611	0.290	0.052
10	173	0.610	0.530	0.052
50	178	0.595	2.145	0.052
100	192	0.581	3.927	0.052
500	337	0.533	15.928	0.052

Todos esses parâmetros variam em função da frequência e das características dos cabos, como seja o diâmetro dos condutores, o tipo de dieléctrico usado, etc. Na tabela 3.1 apresenta-se a variação desses parâmetros em função da frequência, para um cabo com isolamento de polietileno de calibre 0.5 mm, designado segundo a terminologia da *American Wire Gauge* por 24 AWG. Como se pode ver, a capacidade é independente da frequência na banda de frequências consideradas, a indutância tem um decrescimento lento com a frequência, enquanto a resistência e a condutância crescem ambas com a frequência. A resistência para as altas-frequências é proporcional à raiz quadrada da frequência, devido ao efeito pelicular (tendência para a corrente se concentrar na superfície do condutor quando a frequência aumenta).

Os parâmetros secundários de uma linha, definidos usualmente, são a **impedância** característica Z_0, e a **constante de propagação** γ. Esses parâmetros são calculados directamente em termos dos primários do seguinte modo:

$$Z_0(\omega) = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \tag{3.1}$$

$$\gamma(\omega) = \alpha(\omega) + j\beta(\omega) = \sqrt{(R + j\omega L)(G + j\omega C)} \tag{3.2}$$

onde $\omega = 2\pi f$, f é a frequência, α é o **coeficiente de atenuação** expresso em Neper por quilómetro (Np/km) e β é o **coeficiente de fase** expresso em radianos por quilómetro (rad/km). Note-se que, para obter o coeficiente de atenuação em dB/km é necessário um factor multiplicativo de $20\log e \cong 8.68 \text{ dB/Np}$.

Uma linha diz-se que está adaptada se for terminada por uma carga igual à sua impedância característica. Uma linha de comprimento l, adaptada nas duas extremidades, apresenta uma função de transferência dada por

$$H_l(f,l) = \frac{V_r(f,l)}{V_e(f,0)} = e^{-l\gamma(f)} = e^{-l\alpha(f)}e^{-jl\beta(f)} \tag{3.3}$$

onde $V_e(f,0)$ e $V_r(f,l)$ é a transformada de Fourier da tensão, respectivamente no início e no fim da linha. A partir de (3.3) pode-se calcular a **atenuação da linha** para a frequência f, resultando

$$A_{dB}(f,l) = -20\log\left|H(f,l)\right| \cong 8.686\, l\,\alpha(f) \tag{3.4}$$ [27]

Como se pode concluir através de (3.4) a atenuação de uma linha aumenta directamente com o comprimento da linha, sendo também uma função crescente da frequência como é aparente a partir de (3.2).

Outra característica importante dos meios de transmissão é o **atraso de grupo** (expresso em s/km), que não é mais do que o inverso da velocidade de grupo e, por conseguinte, dado por:

$$\tau_g = \frac{d\beta(f)}{2\pi df} \tag{3.5}$$

Quando o atraso de grupo de um determinado meio é uma constante isto significa que esse meio introduz um atraso de propagação constante para todas as frequências presentes no espectro de Fourier do sinal transmitido. Em alternativa, quanto o atraso de grupo depende da frequência, diferentes componentes espectrais têm diferentes tempos de propagação, o que conduz a uma distorção do sinal, designada por **distorção de fase.** A dependência do coeficiente de atenuação com a frequência conduz à distorção de amplitude. Um meio de transmissão que introduz distorção de fase e/ou de amplitude designa-se por meio dispersivo.

3.1.1.2 Aproximações para os parâmetros da linha

Normalmente, no sentido de ganhar uma compreensão mais detalhada das propriedades físicas das linhas de transmissão, recorre-se a determinadas aproximações para as equações (3.1) e (3.2). Uma dessas aproximações consiste em considerar o caso em que a reactância indutiva da linha, assim como a condutância, são desprezáveis, o que corresponde às frequências para as quais se verifica a seguinte relação:

$$\omega L << R \text{ e } G << \omega C \tag{3.6}$$

Nesta banda de frequências tem-se que:

---

[27] Note que se usa o factor multiplicativo de 20 porque se está a trabalhar com tensões.

$$Z_0 \cong \sqrt{\frac{R}{j\omega C}} = \sqrt{\frac{R}{\omega C}}\, e^{-j\frac{\pi}{4}} \tag{3.7}$$

$$\gamma = \alpha + j\beta \cong \sqrt{\frac{\omega RC}{2}} + j\sqrt{\frac{\omega RC}{2}} \tag{3.8}$$

Estas expressões permitem concluir que quando a condição (3.6) é verificada a impedância característica da linha é complexa e varia proporcionalmente a $1/\sqrt{f}$. Por sua vez, o coeficiente de atenuação e de fase são iguais e proporcionais a $\sqrt{f}$, o que implica a existência de distorção de amplitude e de fase. Outra aproximação consiste em considerar as frequências para as quais

$$\omega L >> R \quad \text{e} \quad \omega \text{C} >> \text{G} \tag{3.9}$$

Nesta banda de frequências tem-se que:

$$Z_0 \cong \sqrt{\frac{L}{C}} \tag{3.10}$$

$$\alpha \cong \frac{R}{2}\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}} \tag{3.11}$$

$$\beta \cong \omega\sqrt{LC} \tag{3.12}$$

As equações anteriores mostram que para as frequências que satisfazem (3.9) a impedância característica é real e independente da frequência, e que a constante de fase cresce linearmente com a frequência, o que conduz a um atraso de grupo constante, ou seja, não há distorção de fase. Por sua vez, o coeficiente de atenuação é proporcional a *R*, e só pode considerar-se independente da frequência quando o efeito pelicular for desprezável. Caso contrário, a variação de *R* com $\sqrt{f}$ (como se intui da Tabela 2.1) faz com que o coeficiente de atenuação também seja proporcional a $\sqrt{f}$.

Em síntese, o coeficiente de atenuação apresenta uma variação com $\sqrt{f}$, tanto nas baixas como nas altas frequência. Na região de transição, o crescimento com a frequência é mais lento, sendo aproximadamente proporcional a $f^{1/4}$(ver Figura 3.2). Por sua vez, o atraso de grupo decresce com a frequência, atingindo um valor constante para as altas-frequências. Para frequências muito baixas, o atraso de grupo tem uma variação muito rápida, conduzindo a distorções de fase muito elevadas (ver Figura 3.3).

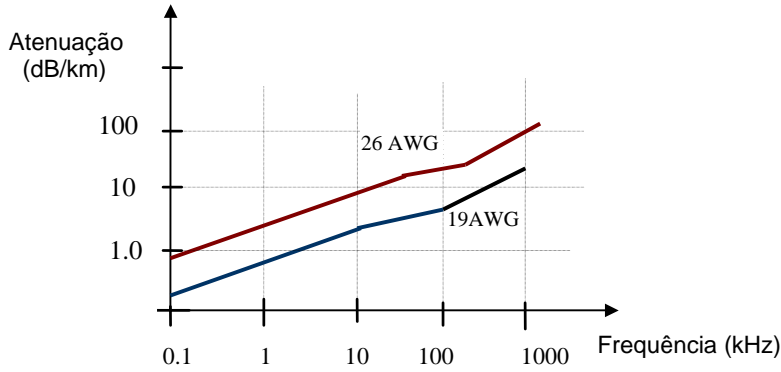



Figura 3.2 Variação da atenuação em função da frequência.

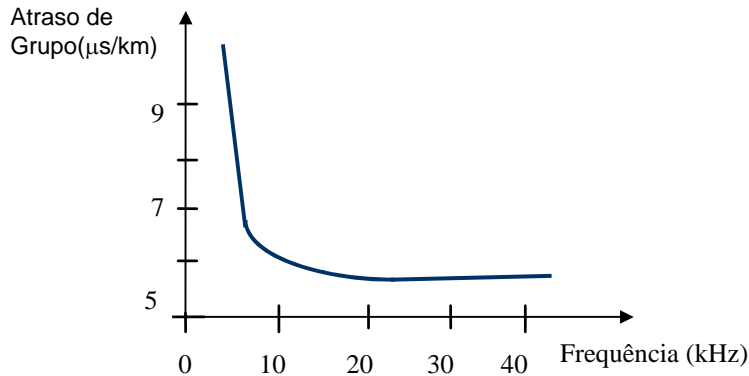



Figura 3.3 Variação do atraso de grupo em função da frequência para um cabo 24 AWG.

O comportamento da atenuação em função da frequência descrito anteriormente leva a considerar-se em muitos casos a seguinte aproximação:

$$\alpha(f) = \alpha_0 \sqrt{f / f_0} \qquad (3.13)$$

onde α_0 é o coeficiente de atenuação à frequência f_0.

As equações (3.11) e (3.12) mostram que no caso em que a reactância indutiva é maior do que a resistência e o efeito pelicular não é significativo a linha tem um comportamento aproximadamente linear, ou seja, não introduz nem distorção de amplitude, nem distorção de fase.

No fim do séc. XIX Pupin teve a ideia de reproduzir essas condições nas baixas frequências, introduzindo na linha bobinas discretas em intervalos regulares. Daí a designação de linhas pupinizadas, ou linhas carregadas. A presença dessas indutâncias vai fazer com que a atenuação se mantenha sensivelmente constante até se atingir a frequência de corte, a partir da qual a atenuação cresce bruscamente (ver Figura 3.4). Essa frequência de corte pode ser calculada usando a equação:

$$f_c = \frac{1}{\pi\sqrt{L_p C d_p}} \qquad (3.14)$$

onde L_p é o valor de indutância das bobinas de carga e d_p é a distância entre as bobinas.

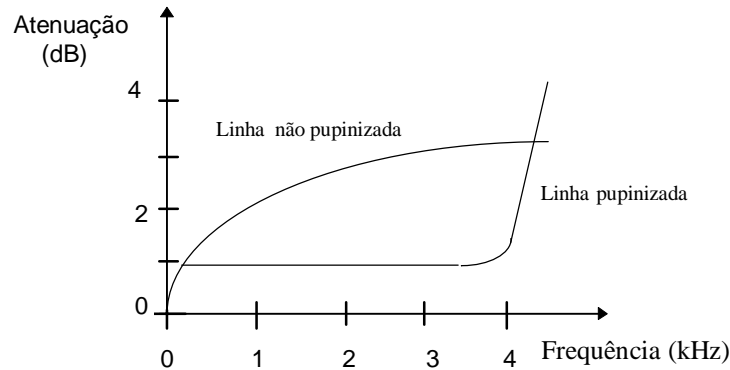



Figura 3.4 Comparação da atenuação de uma linha pupinizada com a de uma não-pupinizada.

Escolhendo, convenientemente, o valor da indutância das bobinas de carga e do espaçamento entre estas, é possível garantir que na banda necessária para a transmissão de voz analógica a atenuação dos pares simétricos seja aproximadamente constante. Esta solução é particularmente interessante quando os utilizadores estão muitos afastados da central local, como acontece nas zonas rurais. No entanto, será importante realçar, que os pares carregados não podem ser usados para a transmissão da voz digitalizada e, por conseguinte, não podem servir de infra-estrutura de transmissão nos lacetes digitais de assinante.

Será de notar que os pares simétricos também são muito usados nas redes de dados, particularmente nas LANs. Para estas aplicações estão disponíveis os pares não blindados ou UPT (*unshielded twisted pair*) e os pares blindados. Os primeiros são os mais usados particularmente para débitos até 10 Mbit/s. Quando os débitos em jogo e o número de pares num cabo são elevados é mais aconselhável usar os segundos, pois a blindagem permite reduzir as interferências electromagnéticas entre pares e, assim, reduzir os efeitos perniciosos da diafonia, os quais se vão estudar na subsecção seguinte.

3.1.1.3 Diafonia

Como já se referiu os pares simétricos não são mais do que dois fios de cobre isolados e entrelaçados. Os cabos de pares simétricos podem, por exemplo, ser construídos agregando vários grupos de pares simétricos, como se mostra na Figura 3.5.

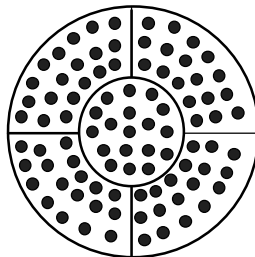



Figura 3.5 Cabo de 90 pares simétricos estruturado em 5 grupos.

A proximidade dos pares no cabo vai originar interferências mútuas entre os diferentes pares, designadas por **diafonia** ou ***crosstalk.*** Estas interferências têm origem, fundamentalmente, no acoplamento capacitivo entre os condutores de dois pares, assim como, no acoplamento indutivo resultante do campo magnético de um dos pares atravessar o campo dos outros pares.

A diafonia é uma limitação séria dos cabos de pares simétricos, principalmente, no que diz respeito à transmissão digital. Podem-se identificar dois tipos de diafonia: a **paradiafonia** e a **telediafonia**. Estes tipos são ilustrados conceptualmente na Figura 3.6. Nessa figura admite-se que o sinal $v_s(t)$ é aplicado à entrada do par perturbador. À medida que este sinal se propaga vai induzir dois tipos de sinais interferentes no par adjacente, designado por par perturbado. O sinal $v_p(t)$, que é induzido na extremidade esquerda, e corresponde à paradiafonia[28] e o sinal $v_t(t)$, que surge na extremidade direita, e representa a telediafonia[29]. A paradiafonia pode constituir uma limitação séria, já que o seu nível pode ser da mesma ordem de grandeza do sinal recebido proveniente da extremidade D. Essa situação ocorre por exemplo, no caso da transmissão ADSL (ver capítulo 5), quando dois utilizadores usam pares simétricos que partilham o mesmo cabo e o modem de um utilizador está em processo de emissão e o modem do outro utilizador está em processo de recepção.

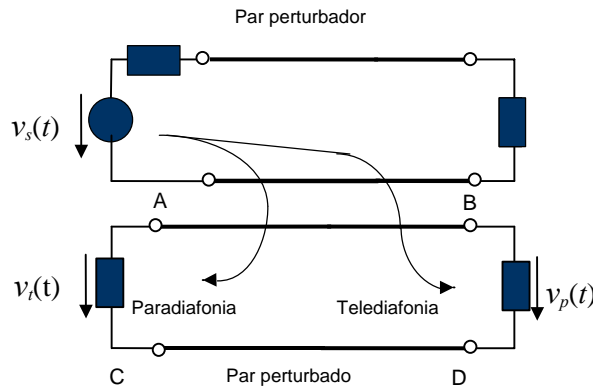



Figura 3.6 Tipos de diafonia nos pares simétricos.

Para calcular a relação sinal-diafonia é necessário conhecer a densidade espectral de potência da diafonia. Admitindo que a densidade espectral de potência do sinal $v_s(t)$é dada por $S(f)$, pode-se escrever a densidade espectral de potência de $v_p(t)$ (paradiafonia) originada por uma multiplicidade de pares como

$$S_p(f) = S(f)\left|X_p(f)\right|^2 \cong S(f)\chi_p f^{3/2} \tag{3.15}$$

onde $X_p(f)$ é a função de transferência da paradiafonia e χ_p é um parâmetro que depende das características do par perturbado. Para um par com isolamento de polietileno

[28] Também designada por NEXT (*Near-end-crosstalk)*
[29] Também designada por FEXT (*Far-end-crosstalk*)

e calibre de 0.5 mm (24 AWG), um valor típico para esse parâmetro é de 1.7x10^{-9} kHz$^{-3/2}$. Para o caso da telediafonia, a densidade espectral de potência pode ser dada por

$$S_t(f) = S(f)\left|X_t(f)\right|^2 \cong S(f)\chi_t f^2 l e^{-2\alpha(f)l} \tag{3.16}$$

onde $X_t(f)$ é a função de transferência da telediafonia, l é o comprimento do cabo, $\alpha(f)$ é o coeficiente de atenuação e χ_t é um parâmetro cujo valor típico é igual a 10^{-10} kHz^{-2}km^{-1} para o par com as características referidas.

Como mostram as equações (3.15) e (3.16) o efeito dos dois tipos de diafonia cresce com o aumento da frequência. Assim, será de esperar que, no caso dos lacetes de assinante analógicos, que usam a banda de transmissão situada entre os 300 e os 3400 Hz a sua acção seja desprezável. Porém, nos lacetes digitais, que podem usar bandas da ordem das várias centenas de kHz[30], o efeito da diafonia é muito problemático e não pode ser ignorado pelo projectista destes lacetes.

3.1.2 Cabos de pares coaxiais

Um par coaxial é constituído por dois condutores concêntricos, sendo o condutor interior isolado e centrado com ajuda de um material dieléctrico (ver Figuras 3.6 e 3.7). Os pares coaxiais são agrupados em pequeno número num cabo protegido por uma bainha com propriedades mecânicas apropriadas para resistir aos esforços de tracção colocados.

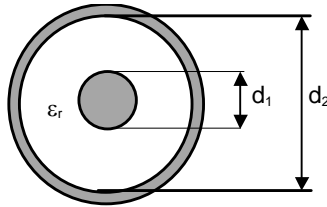



Figura 3.6 Estrutura de um par coaxial.

A impedância característica de um par coaxial, dentro da hipótese que $\omega L >> R$ (largamente verificada no domínio de utilização destes cabos) é dada por

$$Z_o = \frac{60}{\sqrt{\varepsilon_r}} \ln\frac{d_2}{d_1} \tag{3.17}$$

onde ε_r é a permitividade relativa do dieléctrico (=1.074) e d_1 e d_2 são, respectivamente, os diâmetros dos condutores interior e exterior[31]. O coeficiente de atenuação pode ser calculada com base na expressão

$$\alpha = a + b\sqrt{f} + cf\,, \qquad \text{dB/km} \tag{3.18}$$

[30] Os lacetes que suportam o ADSL requerem cerca de 1 MHz.
[31] Note-se que este diâmetro é medido do lado interior.

onde a, b e c são três constantes, que dependem das dimensões físicas do cabo e f é frequência dada em MHz. Para um cabo de 1.2/4.4 mm têm-se a=0.07, b=5.15 e c=0.005. Verifica-se, facilmente, que o termo dominante na equação (3.18) é o termo com dependência em $\sqrt{f}$, o que permite aproximar a atenuação dos cabos coaxiais por uma expressão similar à (3.13) (notar, no entanto, que 3.18 é expressa em dB/km). Para as frequências de interesse (f>100 kHz) o coeficiente de fase apresenta uma variação linear com a frequência, sendo descrita por

$$\beta = \frac{\sqrt{\varepsilon_r}}{c}\omega \qquad (3.19)$$

onde c é a velocidade de propagação da luz no vazio. Como se conclui o atraso de grupo só depende das características do dieléctrico ($\tau_g = \sqrt{\varepsilon_r} / c$), ou seja, o par coaxial não introduz distorção de fase para as frequências referidas.

Pode-se demonstrar, que a atenuação do cabo é mínima para uma relação entre o diâmetro do condutor exterior e condutor interior igual a 3.6. Por isso, os cabos coaxiais normalizados pelo ITU-T têm uma relação próxima daquele valor, como se pode constatar a partir da Tabela 3.2.

Tabela 3.2: Pares coaxiais normalizados

Tipo	2.6/9.5	1.2/4.4	0.7/2.9
Recomendação ITU-T	G 623	G 622	G 621
d_1	2.6 mm	1.2 mm	0.7 mm
d_2	9.5 mm	4.4 mm	2.9 mm
d_2/d_1	3.65	3.67	4.14
Z_0	75±1 Ω	75±1 Ω	75±1 Ω

A estrutura dos pares coaxiais elimina praticamente a possibilidade de acoplamentos capacitivos e indutivos entre os diferentes pares. Assim, embora possam surgir acoplamentos galvânicos entre os condutores exteriores, as propriedades de diafonia deste meio de transmissão são excelentes para frequências de operação acima dos 60 kHz. Abaixo dos 60 kHz, os cabos coaxiais não devem ser usados devido ao problema da captação de interferências exteriores (diafonia elevada), assim como ao facto de introduzirem distorção de fase.

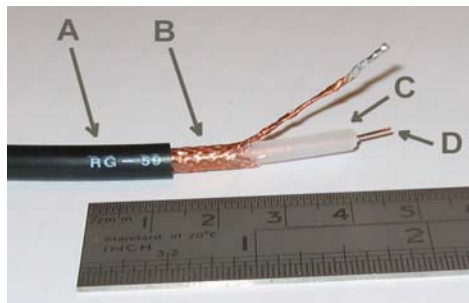



Figura 3.7 Estrutura do cabo coaxial RG-59. A: cobertura protectora de plástico; B: malha de cobre; C: isolador; D: núcleo de cobre.

Os cabos coaxiais constituíram o meio de transmissão por excelência dos sistemas de transmissão analógica de grande distância, tendo permitido suportar sistemas FDM de grande capacidade, como sejam os sistemas definidos pela recomendação G.333 do ITU-T (10 800 canais telefónicos, com uma frequência máxima de 60 MHz). Nos dias de hoje, o cabo coaxial praticamente deixou de ser usado nas redes telefónicas, tendo o seu lugar sido ocupado pelas fibras ópticas. Ocupa, contudo, um lugar importante nas redes locais de computadores (LAN) e nas redes híbridas (fibra/coaxial) das redes de distribuição de televisão por cabo.

3.1.3 Fibras ópticas

Desde o fim dos anos setenta que as fibras ópticas se tornaram um dos meios de transmissão mais importantes para os sistemas de telecomunicações de média e longa distância, tendo vindo também a ganhar alguma relevância na curta distância. Em Portugal, uma grande percentagem do tráfego das redes de transporte é transportado usando meios ópticos. Apresenta-se, em seguida, algumas das razões que explicam o porquê do enorme sucesso das fibras ópticas:

- **Baixa atenuação:** as perdas de transmissão introduzidas pelas fibras ópticas são muito reduzidas quando comparadas, quer com os pares simétricos, quer com os cabos coaxiais (ver Figura 3.8).

- **Largura de banda elevada**: a fibra óptica tem capacidade para transmitir sinais de frequências muito mais elevadas do que o cabo coaxial. A largura de banda de transmissão disponível também depende do tipo de fibra, sendo a fibra monomodal a que apresenta maior capacidade (cerca de 25 000 GHz só na terceira janela de transmissão).

- **Dimensões e peso reduzidos**: a fibra óptica é mais leve e tem um diâmetro mais reduzido do que qualquer outro meio de transmissão metálico. Comparando um cabo coaxial com 18 pares coaxiais, com um cabo óptico com 18 fibras, o cabo óptico ocupa uma secção que é 1/10 da do cabo coaxial e o seu peso é cerca de 1/30. Assim, os condutas usados pelas empresas de telecomunicações podem acomodar cerca de 10 cabos ópticos, no mesmo espaço onde acomodavam um cabo coaxial.

- **Imunidade a interferências electromagnéticas**: como o material base das fibras ópticas é o vidro de sílica (SiO2) e este não conduz electricidade, a fibra óptica é imune às interferências electromagnéticas induzidas por fontes exteriores (Ex. cabos de alta tensão, radiodifusão, descargas atmosféricas, etc.), como também é imune à diafonia originada pela presença de outras fibras.

- **Custo reduzido**: o cobre é um recurso muito limitado, enquanto a matéria-prima usada no fabrico das fibras de vidro (sílica, SiO_2) é um dos materiais mais abundantes na natureza. Actualmente, as fibras ópticas já são mais baratas do que os meios de cobre. Além disso, como a atenuação da fibra, assim como a dispersão, são muito inferiores às dos meios de cobre, os sistemas de transmissão óptica requerem um número muito

menor de repetidores, conduzindo a uma redução muito significativa do preço total do sistema.

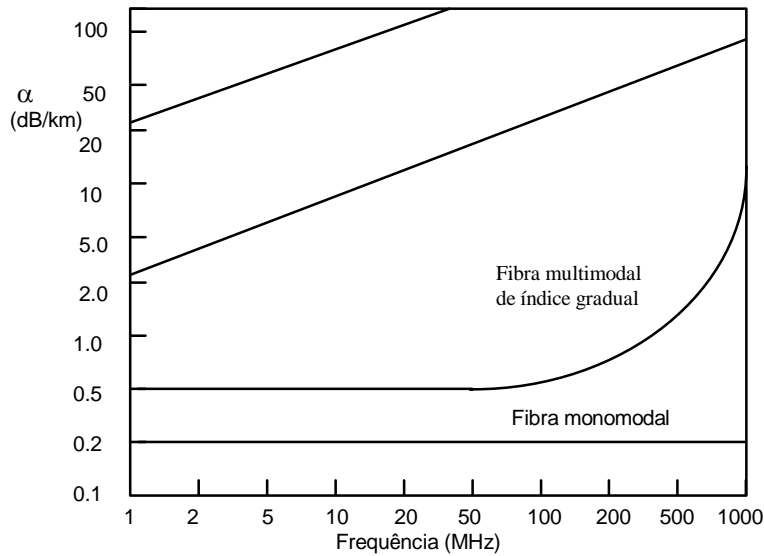



Figura. 3.8 Comparação da atenuação entre os meios metálicos e as fibras ópticas.

Fundamentalmente, uma fibra óptica é um guia dieléctrico cilíndrico constituído por um núcleo e por uma bainha, sendo o índice de refracção do núcleo (n_1) superior ao índice de refracção da bainha (n_2) (ver Figura 3.9).

As fibras ópticas dividem-se em dois grupos: as fibras **multimodais** e as fibras **monomodais**. Como se sabe dos estudos de propagação, realizados em disciplinas anteriores, as primeiras suportam vários modos de propagação, enquanto as últimas suportam um único modo (para ser mais rigoroso dois modos ortogonais). As fibras monomodais apresentam um diâmetro do núcleo de poucos micrómetros (2 a 10 μm). As fibras multimodais são de mais fácil manuseamento, porque o diâmetro do núcleo é maior (entre 25 e 100 μm)[32], mas em alternativa apresentam uma dispersão mais elevada, o que conduz a uma largura de banda menor.

As principais características de transmissão das fibras ópticas são a **atenuação** e a **dispersão.** A atenuação total apresentada por uma fibra óptica é dada por

$$A(\text{dB}) = \alpha l \tag{3.20}$$

onde l é o comprimento da fibra (em km) e α é o coeficiente de atenuação, em dB/km. O valor do coeficiente de atenuação depende do comprimento de onda de trabalho. Os sistemas de transmissão óptica presentes operam em duas bandas de comprimentos de onda, também designadas por **janelas de transmissão**: a janela de 1.3 μm e a janela de 1.55 μm, as quais pertencem ambas ao domínio do infravermelho. A atenuação das fibras ópticas é mínima para a janela de 1.55 μm, apresentando um valor típico de 0.2 dB/km.

[32] Os valores típicos são de 50 e 62.5 μm.

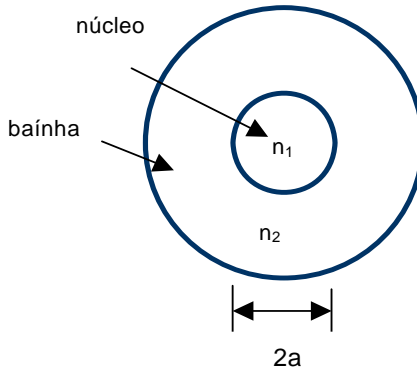



Figura 3.9 Estrutura simplificada de uma fibra óptica.

A dispersão, como se sabe, divide-se em **intermodal** e **intramodal.** A primeira resulta do facto de diferentes modos apresentarem diferentes tempos de propagação, enquanto a segunda se deve ao facto de diferentes comprimentos de onda terem diferentes tempos de propagação. As fibras ópticas monomodais só apresentam dispersão intramodal, enquanto as fibras multimodais estão sujeitas às duas contribuições.

Antes de se estudar as implicações da dispersão nos sistemas de transmissão, será conveniente fazer uma breve referência às fontes emissoras usadas na transmissão óptica. A fonte emissora usada, na generalidade dos casos, é o laser de semicondutor, também designado por díodo laser. A modulação deste dispositivo é feita variando a corrente de polarização aos seus terminais (modulação directa). Na Figura 3.10 exemplifica-se essa modulação para um sistema de transmissão digital. Como se vê, a intensidade luminosa (potência óptica) varia directamente em resposta à variação da corrente. Assim, pode-se concluir que num sistema de transmissão óptica digital, o transporte da informação está associado a variações da intensidade luminosa.

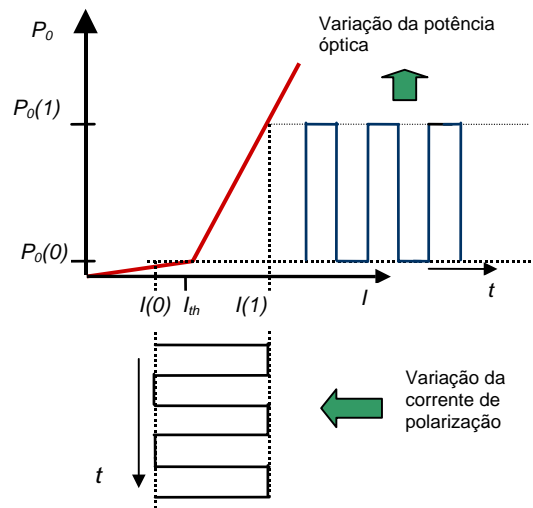



Figura 3.10 Modulação em intensidade de um laser de semicondutor.

Um dos maiores inconvenientes dos díodos laser é que estes não são fontes monocromáticas. Em contrapartida, o seu espectro de emissão (sem modulação) apresenta

uma largura espectral σ_λ, que pode ir de alguns nanómetros, para os laser de menor qualidade (Fabry-Perot), até valores da ordem da milésima do nanómetro, para lasers de melhor qualidade (DFB[33]).

A dispersão intramodal é consequência imediata da falta de coerência espectral das fontes ópticas. Admitindo que se tem uma fibra monomodal, caracterizada pelo parâmetro de dispersão D_λ [ps/nm/km], pode-se calcular de modo imediato o alargamento do pulso óptico devido à dispersão intramodal, σ_t, desde que se conheça o comprimento da fibra l e a largura espectral da fonte óptica usada σ_λ, vindo

$$\sigma_t = |D_\lambda| l \sigma_\lambda \tag{3.21}$$

O conhecimento da função de transferência dos meios de transmissão é de grande importância para um engenheiro de telecomunicações. Uma boa aproximação para a função de transferência de uma fibra óptica monomodal é a seguinte:

$$H(f) = H(0)\exp\left[-2\pi^2 f^2 \sigma_t^2\right] \tag{3.22}$$

A partir da equação anterior pode-se determinar a largura de banda óptica da fibra definida a -3 dB, obtendo-se

$$B_0 = \frac{0.187}{\sigma_t} \tag{3.23}$$

Exemplo 3.1 Pretende-se deduzir uma expressão que permita calcular o débito binário máximo suportado por uma fibra óptica monomodal.

Solução: Como se irá ver na secção (3.44) a largura de banda mínima de um canal que permite suportar um débito binário D_b é dada por

$$B_e \geq D_b / 2$$

Por sua vez, a largura de banda eléctrica (B_e) relaciona-se com a óptica (B_0) através de

$$B_o = B_e\sqrt{2}$$

Atendendo às relações anteriores e à equação (3.23) deduz-se que

$$D_b \leq 0.264 / \sigma_t$$

[33] Distributed Feedback

3.2 *Amplificação versus regeneração*

3.2.1 Transmissão conforme e distorção

Um meio de transmissão conforme é aquele que garante uma transmissão sem deformação do sinal. Um meio de transmissão diz-se conforme ou não distorcivo quando o sinal recebido $v_r(t)$ difere do sinal emitido $v_e(t)$ unicamente por:

- um factor a constante;
- um atraso τ constante.

Nessas circunstâncias, os dois sinais referidos relacionam-se do seguinte modo:

$$v_r(t) = a v_e(t - \tau) \tag{3.24}$$

A esta identidade no domínio temporal, corresponde a seguinte identidade no domínio frequencial

$$V_r(f) = a V_e(f) \exp(-j2\pi f\tau) \tag{3.25}$$

de onde resulta que:

- a atenuação A da transmissão deverá ser constante e independente da frequência:

$$A = 20\log\left|\frac{V_e}{V_r}\right| = 20\log\frac{1}{a} \tag{3.26}$$

- a desfasagem ϕ deve ser uma função *linear* da frequência:

$$\phi = \arg\left[V_e(f)/V_r(f)\right] = -2\pi f\tau \pm m\pi \tag{3.27}$$

onde m indica um número inteiro As condições (3.26) e (3.27) podem não ser satisfeitas nos sistemas práticos. Como já se viu, os meios de transmissão estudados introduzem dois tipos de distorção linear, a distorção de amplitude, reflectida no facto de atenuação A depender da frequência, ou seja, de acordo com (3.4)

$$A(f,l)_{dB} = l\alpha(f)_{dB/km} \tag{3.28}$$

e a distorção de fase resultante do facto de a variação da desfasagem com a frequência não ser linear. A desfasagem relaciona-se com o coeficiente de fase através de

$$\phi(f,l) = \beta(f)l \tag{3.29}$$

e, como se viu, a variação de β com a frequência nem sempre é linear (ver por exemplo a equação (3.8).

Como consequência das distorções lineares, um sinal puramente sinusoidal na emissão permanece sinusoidal na recepção, mas qualquer outro tipo de sinal verá o seu espectro alterado. Contudo, como não são geradas novas componentes espectrais, o princípio da sobreposição permanece válido, daí a designação de distorção linear.

3.2.2 Amplificação

3.2.2.1 Caracterização do ganho

A amplificação tem como objectivo compensar a atenuação introduzida pelo meio de transmissão, de modo a obter um nível do sinal na recepção capaz de garantir as exigências de qualidade requeridas.

Cada amplificador é caracterizado pelo seu ganho, o qual relaciona a potência do sinal na saída p_0, com a potência do sinal na entrada p_i, isto é

$$g = \frac{p_0}{p_i} \tag{3.30}$$

Normalmente, o ganho dos amplificadores usados nos sistemas de telecomunicações é muito elevado, sendo conveniente exprimir o ganho de potência em decibel (dB), que como se sabe corresponde a fazer

$$G = 10\log_{10} g \tag{3.31}$$

Em sistemas de telecomunicações também é usual exprimir-se as potências em unidades logarítmicas, usando como nível de referência um sinal com a potência de 1 W, ou de 1mW, tendo-se, respectivamente,

$$P_{d\mathrm{BW}} = 10\log_{10}\frac{p}{1\,\mathrm{W}} \qquad\qquad P_{\mathrm{dBm}} = 10\log_{10}\frac{p}{1\,\mathrm{mW}} \tag{3.32}$$

Trabalhando com unidades logarítmicas, a equação (3.30) reescreve-se na forma:

$$P_{0\ \mathrm{dBm}} = P_{i\ \mathrm{dBm}} + G_{\mathrm{dB}} \tag{3.33}$$

Neste estudo admite-se que o amplificador apresenta um ganho constante, g_0, em toda a banda de amplificação, como se ilustra na Figura 3.11. A largura de banda a -3dB é igual a B. Tal amplificador não tem capacidade para compensar, nem a distorção de amplitude, nem a distorção de fase. Assim, o efeito da compensação da atenuação obtido com um amplificador com as características referidas só é realmente efectivo, quando o meio de transmissão for conforme. Em todos os outros casos é necessário usar um dispositivo apropriado para compensar as distorções referidas, que se designa por **igualador.**

Como exemplo de uma rede em que o meio de transmissão pode ser considerado aproximadamente conforme pode-se referir a rede híbrida usada como infra-estrutura de

suporte na distribuição de televisão por cabo. O meio de transmissão da parte eléctrica é o cabo coaxial, cuja atenuação se pode considerar aproximadamente constante na banda de transmissão usada.

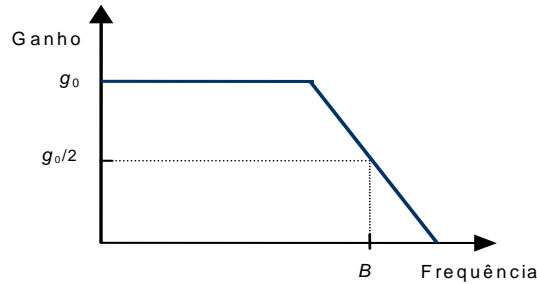



Figura 3.11 Característica de transferência de um amplificador.

3.2.2.2 Ruído

Qualquer amplificador introduz ruído num sistema de transmissão. Admitindo que o amplificador é perfeitamente linear este ruído tem a sua origem em duas fontes:

- Agitação térmica na resistência de entrada do amplificador (ruído termico);
- Ruído electrónico ou interno gerado pelos elementos activos (transistores) no interior do amplificador.

O ruído térmico, como é conhecido, é um ruído branco e Gaussiano, com densidade espectral de potência unilateral dada por

$$S_n(f) = kT \tag{3.34}$$

onde k é a constante de Boltzman (1.38 x 10^{-23} J/ºK) e T é a temperatura ambiente em K. A potência de ruído equivalente na entrada do amplificador é facilmente obtida a partir da densidade espectral de potência resultando em

$$n_i = \int_0^B S_n(f)df = kTB_n \tag{3.35}$$

onde B_n é a largura de banda equivalente de ruído. Designando por s_i a potência do sinal na entrada do amplificador, a relação sinal-ruído na entrada, expressa em unidades logarítmicas, pode-se escrever na forma:

$$\left(\frac{S_i}{N_i}\right)_{d\mathrm{B}} = 10\log_{10}\left(\frac{s_i}{n_i}\right) = 10\log_{10}\left(\frac{s_i}{kTB_n}\right) \tag{3.36}$$

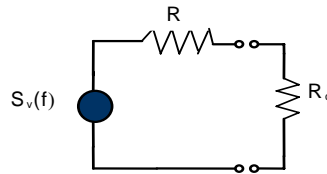
A segunda fonte de ruído está localizada nos elementos activos do amplificador e constitui o seu ruído interno. Como consequência, a relação sinal-ruído na saída do

amplificador é inferior à relação sinal-ruído na entrada, calculada usando (3.36). Define-se o **factor de ruído** do amplificador (em dB), como a diferença entre a relação sinal-ruído na entrada (em dB) e a relação sinal-ruído na saída (em dB), ou seja

$$F_{\mathrm{dB}} = \left(\frac{S_i}{N_i}\right)_{\mathrm{dB}} - \left(\frac{S_0}{N_0}\right)_{\mathrm{dB}} \tag{3.37}$$

Exemplo 3.1

Pretende-se demonstrar que a densidade espectral de potência do ruído térmico é dada pela equação (3.34).

Solução: Para isso deve-se atender a que o circuito equivalente de uma resistência a uma determinada temperatura T, se pode representar pelo seguinte esquema:

Esse esquema representa uma resistência R ideal sem ruído em série com uma fonte de tensão de ruído com densidade espectral de potência de dada por

$$S_v(f) = 4RkT \ \ \mathrm{V^2Hz^{-1}}$$

Por sua vez, a densidade espectral de potência, correspondente à potência disponível aos terminais da resistência de carga R_c, vem dada por

$$S_n(f) = \frac{S_v(f)R_c^2}{R_c(R_c + R)^2}$$

No caso específico em que a carga está adaptada ($R_c = R$), as equações anteriores conduzem a

$$S_n(f) = \frac{S_v(f)}{4R} = kT$$

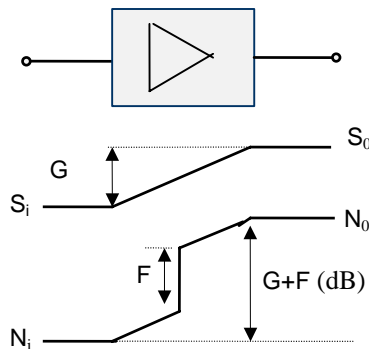



Figura 3.12 Degradação da relação sinal-ruído num amplificador.
(Na figura usam-se unidades logarítmicas)

3.2.2.3 Efeito da não linearidade

Num amplificador ideal a tensão instantânea de saída $v_o(t)$ deve ser exactamente proporcional à tensão de entrada $v_i(t)$. Os amplificadores reais aproximam-se da idealidade unicamente num domínio de tensão limitado (ver Figura 3.13). Fora desse domínio a tensão de saída $v_o(t)$ deixa de ser proporcional à tensão de entrada, tendendo para um valor constante (tensão de saturação).

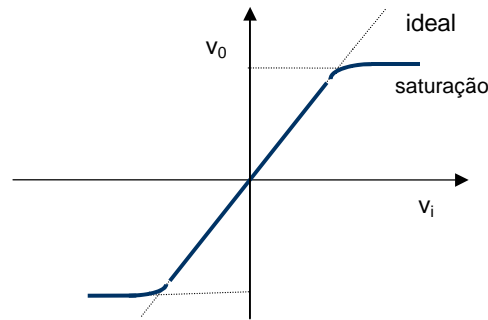



Figura 3.13 Característica de um amplificador real.

Quando o amplificador opera em saturação o sinal é distorcido. No caso específico de um sinal de entrada sinusoidal são originadas harmónicas em outras frequências, designadas por **produtos de intermodulação.** Um sinal aleatório qualquer (voz, vídeo, etc.) fica sujeito a uma degradação adicional, designada por ruído de intermodulação, cuja influência aumenta, quando o nível do sinal de entrada aumenta. Nas redes híbridas de distribuição de televisão este ruído de intermodulação é um factor limitativo importante, que tem de ser obviamente considerado no seu projecto. De notar que o comportamento descrito só é válido para os amplificadores que operam no domínio eléctrico. Os amplificadores ópticos designados por EDFAs[34], que são usados nos sistemas de transmissão óptica, operam normalmente em saturação, sem introduzirem nenhuma distorção significativa no sinal amplificado.

3.2.3 Sistema de transmissão com repetidores não regenerativos

Considere-se um sistema de transmissão constituído por m troços de cabo, cada um terminado por um amplificador, designado por repetidor não regenerativo, como se mostra na Figura 3.14. Admite-se que as impedâncias de entrada e de saída de cada amplificador estão adaptadas ao cabo. Designe-se por $g_{c,m}$ e $g_{a,m}$, respectivamente, o ganho do cabo do troço m e o ganho do amplificador m. Note-se que $g_{c,m}<1$. Então o conjunto constituído pelo troço de cabo m e pelo amplificador m apresenta um ganho dado por

$$g_m = g_{c,m}\ g_{a,m} \tag{3.38}$$

[34] *Erbium Doped Fiber Amplifier*

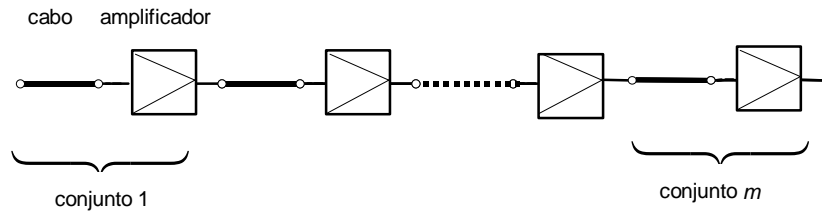



Figura 3.14 Configuração de um sistema de transmissão com *m* amplificadores.

Será útil conhecer também o factor de ruído do conjunto referido. Para isso, deve atender-se que, se o cabo estiver adaptado e em equilíbrio com o meio ambiente, a potência de ruído à saída é igual à potência de ruído à entrada. Então, pode-se demonstrar que o factor de ruído do conjunto (cabo, amplificador) *m*, apresenta um factor de ruído (linear) dado por

$$f_m = \frac{f_a}{g_{c,m}} \tag{3.39}$$

onde f_a é o factor de ruído do amplificador. A partir do conhecimento do ganho e do factor de ruído de cada um dos *m* conjuntos pode-se caracterizar a cadeia constituída pelos *m* troços de cabo e pelos *m* amplificadores. Assim, o ganho da cadeia é dado por

$$g = g_1\, g_2\, g_3 \,.......g_{\mathrm{m}} \tag{3.40}$$

enquanto o factor de ruído é descrito por

$$f = f_1 + \frac{f_2 - 1}{g_1} + \frac{f_3 - 1}{g_1 g_2} + \cdots\cdots + \frac{f_m - 1}{g_1 g_2 \cdots g_{m-1}} \tag{3.41}$$

Dentro da hipótese simplificativa que o ganho e o factor de ruído dos diferentes conjuntos da cadeia, são iguais, respectivamente, ao ganho e ao factor de ruído do conjunto, as expressões anteriores simplificam-se, conduzindo a

$$g = g_1{}^m \tag{3.42}$$

$$f = 1 + \sum_{k=1}^{m} \frac{f_1 - 1}{g_1{}^{k-1}} = 1 + (f_1 - 1)\frac{1 - g_1{}^{-m}}{1 - g_1{}^{-1}} \tag{3.43}$$

O conhecimento de *g* e de *f* permite relacionar a potência do sinal na saída da cadeia s_o, com a potência do sinal na entrada s_i, assim como a potência do ruído na saída n_o com a potência do ruído na entrada n_i, tendo-se:

$$\frac{s_o}{n_o} = \frac{g\, s_i}{g\, f\, n_i} = \frac{s_i}{f n_i} \tag{3.44}$$

Como se conclui facilmente, a equação anterior traduz uma outra forma de apresentar a equação (3.37).

3.2.4 Transmissão digital e regeneração

3.2.4.1 Caracterização da informação digital

A informação gerada por uma fonte digital é representada, tendo em vista a transmissão, por símbolos, ou sinais elementares de duração T_s. Esses símbolos pertencem a um alfabeto de M símbolos, como sejam por exemplo M níveis de amplitude, ou M fases. A velocidade física de transmissão desses sinais elementares designa-se por débito de símbolo $D_s = 1/T_s$, sendo expressa em baud (bd). Por sua vez a informação é gerada a uma velocidade designada por débito binário D_b (bit/s). Estas duas grandezas estão relacionadas por

$$D_b = D_s \log_2 M \tag{3.45}$$

Como consequência o tempo de bit é dado por

$$T_b = \frac{1}{D_b} = \frac{T_s}{\log_2 M} \tag{3.46}$$

3.2.4.2 Influência do canal

Como se sabe, a partir do conhecimento da função de transferência de um canal de transmissão $H(f)$ e da transformada de Fourier do sinal de entrada $V_e(f)$, determina-se de forma imediata, a transformada de Fourier do sinal de saída (ver por exemplo (3.3)):

$$V_r(f) = V_e(f)H(f) \tag{3.47}$$

As características do canal de transmissão dependem do meio de transmissão usado. Como se viu, os meios de transmissão estudados não são meios conformes, introduzindo por isso distorções de amplitude e de fase.

Se o sinal transmitido for um sinal digital, essa não idealidade do canal de transmissão, vai originar interferência entre os diferentes símbolos, a qual se designa na literatura técnica por **interferência intersimbólica**. A presença de interferência intersimbólica pode conduzir a erros de transmissão, resultantes do facto de a informação recebida não ser exactamente idêntica à informação enviada pela fonte.

3.2.4.3 Igualação

Através de um processamento adequado do sinal na saída de um determinado canal de transmissão é possível minimizar, ou mesmo eliminar, a influência da interferência

intersimbólica. Esse processamento é feito usando um filtro designado por igualador. Para o projectista de um sistema de transmissão digital é fundamental conhecer a função de transferência desse filtro. Esse cálculo pode ser feito atendendo à Figura 3.15

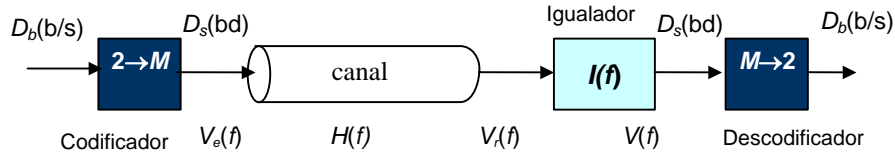



Figura 3.15 Igualação num sistema de transmissão digital.

De acordo com essa figura, a função de transferência do igualador será dada por:

$$I(f) = \frac{V(f)}{V_e(f)\, H(f)} \tag{3.48}$$

onde $V_e(f)$ e $V(f)$ são as transformadas de Fourier do sinal, respectivamente, na entrada do canal e na saída do igualador. Sabe-se dos Fundamentos das Telecomunicações que para garantir uma interferência intersimbólica nula, o sinal na saída do igualador deve satisfazer os critérios de Nyquist. Para tal é necessário que $V(f)$ seja descrito por uma função apropriada, capaz de satisfazer esses critérios. Como exemplo, pode-se referir a função coseno-sobreelevado, tendo-se assim

$$V(f) = \begin{cases} T_s & 0 \le |f| \le f_N(1-\alpha) \\ \frac{T_s}{2}\left[1 - \sin\left(\frac{\pi|f|}{2\alpha f_N} - \frac{\pi}{2\alpha}\right)\right] & f_N(1-\alpha) \le |f| \le f_N(1+\alpha) \\ 0 & |f| \ge f_N(1+\alpha) \end{cases} \tag{3.49}$$

Nessa função α representa o factor de excesso de banda (α pode variar entre 0 e 1) e f_N a frequência de Nyquist que é definida por

$$f_N = \frac{1}{2T_s} \tag{3.50}$$

A equação (3.49) e a Figura 3.16 sugerem que a largura de **banda mínima** de um canal de transmissão, que garante uma comunicação com interferência intersimbólica nula é igual a

$$B_{\min} = f_N = \frac{1}{2T_s} = \frac{D_s}{2} = \frac{D_b}{2\log_2 M} \tag{3.51}$$

No caso concreto de uma transmissão binária, a largura de banda do canal mínima é igual a $D_b/2$. É claro que a função de transferência que conduz à banda mínima não é fisicamente realizável. Por isso, para um determinado débito binário, a largura de banda de transmissão de um canal real é sempre superior à banda mínima.

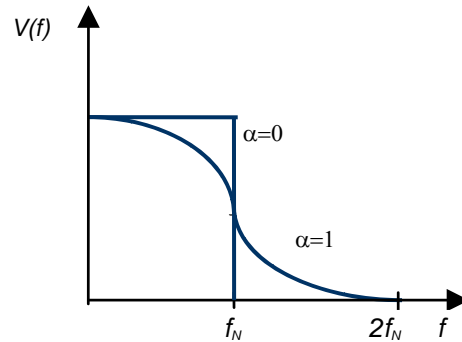



Figura 3.16 Função coseno sobreelevado para dois valores do factor de excesso de banda.

3.2.4.4 Regeneração

A regeneração consiste em reconstituir tão fielmente quanto possível a informação digital transportada por um sinal, depois deste ter sido atenuado, distorcido e perturbado, em consequência da sua propagação num canal de transmissão real.

A informação reconstituída é entregue ao destinatário (regenerador terminal), ou é novamente transmitida pelo canal de transmissão (regenerador intermediário ou repetidor regenerativo).

Na figura 3.17 representa-se a estrutura de um regenerador, para um sistema usando um código binário. As funções principais desse regenerador são as seguintes:

- **igualação**: compensar a distorção introduzida pelo canal de modo a minimizar a interferência intersimbólica;

- **amplificação**: compensar a atenuação introduzida pelo canal, sendo as variações da atenuação do canal - devidas nomeadamente às variações de temperatura - compensadas pelo controlo automático de ganho.

- **recuperação de relógio**: recuperar a frequência de relógio a partir do sinal recebido e gerar uma sequência de pulsos com essa frequência Se o sinal recebido apresenta componentes espectrais discretas (riscas) à frequência de relógio, pode-se recuperar um sinal sinusoidal com um simples filtro ou PLL e usá-lo para gerar o relógio (sequência de pulsos). Caso contrário, é necessário gerar essas riscas através de um processamento não linear adequado.

- **amostragem**: obter amostras do sinal recebido nos instantes iT_s com a frequência de relógio $f_{rel} = D_s = 1/T_s$; recorde-se que no caso da transmissão binária $T_s = T_b$;

- **decisão e formatação**: comparar as amostras com um limiar de decisão (no caso da transmissão binária), decidir qual o símbolo transmitido e, em seguida, dar ao sinal a forma original (quer em amplitude, quer em duração).

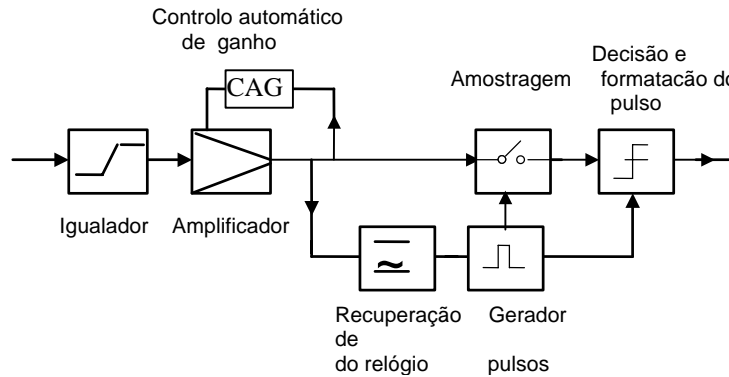



Figura 3.17 Esquema de blocos de um regenerador.

O regenerador descrito também é muitas vezes na literatura técnica com a designação de repetidor 3R[35], porque realiza três funções essenciais: **R**e-formatação (igualação e amplificação); **R**e-temporização (recuperação de relógio); **R**egeneração (decisão e formatação).

Nos sistemas reais, a reconstrução da sequência dos pulsos transmitidos nunca é totalmente perfeita. Há sempre uma probabilidade finita de o circuito de decisão tomar uma decisão errada sobre o símbolo transmitido, e há sempre algum erro (*jitter*) na temporização da sequência transmitida. As fontes de erro são para além da interferência intersimbólica, o ruído térmico, o ruído electrónico dos amplificadores do regenerador, a diafonia, o ruído impulsivo, etc. O ruído impulsivo é devido, nomeadamente, a influências exteriores, tais como descargas atmosféricas, circuitos eléctricos etc.

3.2.4.5 Razão de erros binários

A análise do desempenho de um sistema de transmissão digital é feita em termos da **probabilidade de erro** ou **razão de erros binários.** O ponto de partida para o cálculo da probabilidade erro consiste em determinar a estatística das diferentes fontes de erro. Normalmente, assume-se que essa estatística é Gaussiana, embora essa hipótese possa ser um tanto grosseira, em fenómenos como a interferência intersimbólica e a diafonia. O passo seguinte consiste em calcular a relação sinal-ruído (dentro da hipótese da aproxiamção Gaussiana) à entrada do circuito de decisão, e em seguida deduzir uma expressão adequada para a probalilidade de erro. Se se considerar como exemplo o código AMI, e se admitir que a tensão do sinal tem uma tensão de pico de v volts, tem-se a seguinte expressão para a probabilidade de erro:

---

[35] Seguindo a mesma lógica os amplificadores podem ser designados por repetidores 1R.

$$P_e = \frac{3}{4}\mathrm{erfc}\left(\frac{1}{2}\sqrt{\frac{s}{n}}\right) \tag{3.52}$$

onde erfc(x) é a função complementar de erro, s é a potência média do sinal ($v^2/2$ [36]) e n é a potência média de ruído. A função complementar de erro pode-se aproximar por:

$$\mathrm{erfc}(x) \cong \frac{e^{-x^2}}{x\sqrt{\pi}}, \qquad x > 3 \tag{3.53}$$

A potência de ruído será calculada tendo em conta as diferentes perturbações presentes no sistema. Para exemplificar, considere-se um sistema de transmissão, cujo igualador (com função de transferência $I(f)$) elimina completamente a interferência intersimbólica e cujo desempenho é condicionado pelo ruído térmico e pela diafonia. Nesta situação, a potência de ruído tem duas componentes, ou seja,

$$n = \sigma_t^2 + \sigma_d^2 \tag{3.54}$$

onde σ_t^2 e σ_d^2, representam, respectivamente, a variância devida ao ruído térmico (incluindo o electrónico) e a variância devida à diafonia. A primeira é dada por

$$\sigma_t^2 = \int_0^\infty S_n(f)\,|I(f)|^2\,df \tag{3.55}$$

onde $S_n(f)$ é a densidade espectral de potência de ruído, calculada de acordo com o tratamento apresentado na subsecção 3.2.4.3 e I(f) é a função de transferência do igualador. Relativamente à diafonia, se se admitir que a paradiafonia é dominante, tem-se

$$\sigma_d^2 = \int_0^\infty S_p(f)|I(f)|^2\,df \tag{3.56}$$

onde $S_p(f)$ é a densidade espectral de potência da paradiafonia (ver eq. 3.15).

Até agora fizeram-se algumas considerações, ainda que sumárias, sobre o cálculo da probabilidade de erro em transmissão digital. Mas não é de menor importância, saber-se medir essa probabilidade de erro em sistemas reais. Existem dois caminhos possíveis para realizar essas medidas:

- **teste fora de serviço**: o equipamento de transmissão é simulado, normalmente, por um gerador de sequências pseudo-aleatórias, e a saída do sistema é comparada bit-a-bit com um padrão de referência gerado localmente num detector de erros (ver Figura 3.18). Repare-se que este tipo de medida implica a remoção do tráfego real do sistema.

[36] Assume-se uma impedância unitária.

Por isso este tipo de medidas é mais indicado para teste de produção, testes de instalação e sistemas experimemtais.

- **teste em serviço**: neste caso são monitorizadas continuamente certas propriedades do tráfego real, ou dos códigos de transmissão usados (ex: violação de polaridade no código AMI, no HDB3, ou em outros códigos de linha), tendo em vista a detecção de erros.

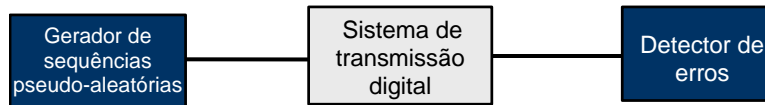



Figura 3.18 Medida da taxa de erros num sistema de transmissão digital.

Uma grandeza normalmente usada neste tipo de medidas é a razão de erros binários ou BER (*bit error ratio*), que se pode obter do seguinte modo:

$$\mathrm{BER} = \frac{N_e}{N_t} \tag{3.57}$$

onde N_e é o número de bits errados contados no intervalo de tempo Δt e N_t é o número total de bits transmitidos no mesmo intervalo de tempo, que é dado por $N_t = D_b \Delta t$.

Os geradores de sequências pseudo-aleatórias são dimensionados para gerarem sequências com um determinado comprimento e de acordo com um determinado polinómio gerador. O comprimento da sequência é determinado pelo número de registos de deslocamento (implementados com *flip-flops D*) usados e o tipo de polinómio pelas ligações de realimentação. Um gerador de sequências com *n* registos, gera uma sequência de comprimento $2^n - 1$. Na Figura 3.19 apresenta-se um exemplo de um gerador de sequências pseudo-aleatórias de comprimento $2^3 - 1$, com o polinómio $1 + D + D^3$. Na Tabela 3.3 apresentam-se outros comprimentos recomendados para diferentes débitos.

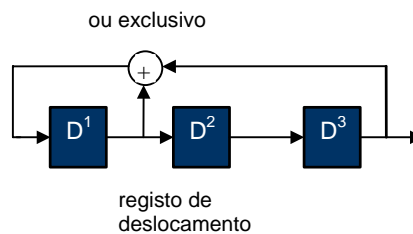



Figura 3.19 Exemplo de um gerador de sequências de comprimento 2^3-1.

Tabela 3.3 Sequências recomendadas pelo ITU-T

Débitos	Comprimento do padrão	Recomendação ITU-T
64 kb/s	2^{11}-1	0.152
2.048 Mb/s	2^{15}-1	0.151
34.368 Mb/s	2^{23}-1	0.151
139.264 Mb/s	2^{23}-1	0.151

3.2.4.6 Sistema de transmissão com repetidores regenerativos em cadeia

Vai-se estudar nesta secção como calcular a probabilidade de erro de um sistema constituído por *m* repetidores regenerativos (regeneradores intermédios) associados em cadeia, considerando que cada repetidor é caracterizado pela probabilidade de erro *p*. Cada dígito transmitido pode estar sujeito a erros cumulativos, à medida que passa de repetidor para repetidor. Se o número total de erros for par, estes cancelam-se. Tem-se um erro no fim da cadeia, só se o número total de erros for ímpar.

Para uma probabilidade *p* de errar um dígito, a probabilidade de originar *k* erros em *m* repetidores é dada pela distribuição binomial:

$$P_k = \binom{m}{k} p^k (1-p)^{m-k} \tag{3.58}$$

A probabilidade de erro da cadeia é obtida somando os valores de P_k, para todos os valores de *k* ímpares, obtendo-se

$$P_e = \sum_{k=1}^{m} \binom{m}{k} p^k (1-p)^{m-k} \qquad k \text{ ímpar} \tag{3.59}$$

Nos sistemas reais em que p«1 e mp«1, essa expressão simplifica-se consideravelmente obtendo-se:

$$P_e \cong mp \tag{3.60}$$

Esta aproximação mostra que, no caso em que todos os regeneradores apresentam a mesma probabilidade de erro e que esta é suficientemente pequena, a probabilidade de erro total cresce linearmente com o número de repetidores.

3.3 Codificação de linha

Em termos gerais os **códigos de linha** têm por objectivo adaptar as características das sequências binárias geradas pelas fontes de informação às características dos suportes de transmissão. As propriedades desejáveis para um código são as seguintes:

- O espectro de potência de um sinal codificado deve apresentar componentes reduzidas nas baixas frequências e deve ser zero à frequência nula;
- O sinal codificado deve apresentar um conteúdo de temporização (transições) suficientemente elevado, de modo a facilitar a recuperação do relógio;
- O código deve ter capacidade para detectar erros, e a sua detecção deve poder ser usada para quantificar o desempenho do sistema (teste em serviço);
- O esquema de codificação deve reduzir a diafonia entre canais.

3.3.1 Tipos de códigos

Na figura 3.18 descrevem-se alguns dos códigos binários mais usuais. As fontes de informação geralmente produzem informação usando um código NRZ (*non-return-to-zero*). Este código tem a particularidade de manter os valores atribuídos aos níveis lógicos "1" e "0", durante todo o intervalo de bit. Os códigos NRZ podem ser unipolares ou polares conforme são assimétricos, ou simétricos relativamente a 0 V. Em termos de potência o NRZ polar é mais eficiente. Como exemplo, considere-se que A=5 V, então ao código unipolar corresponde uma potência de 12.5 W (assumindo uma resistência de 1 Ω), enquanto ao polar corresponde uma potência de 6.25 W (admitiu-se que os níveis do polar variam entre 2.5 e –2.5 V).

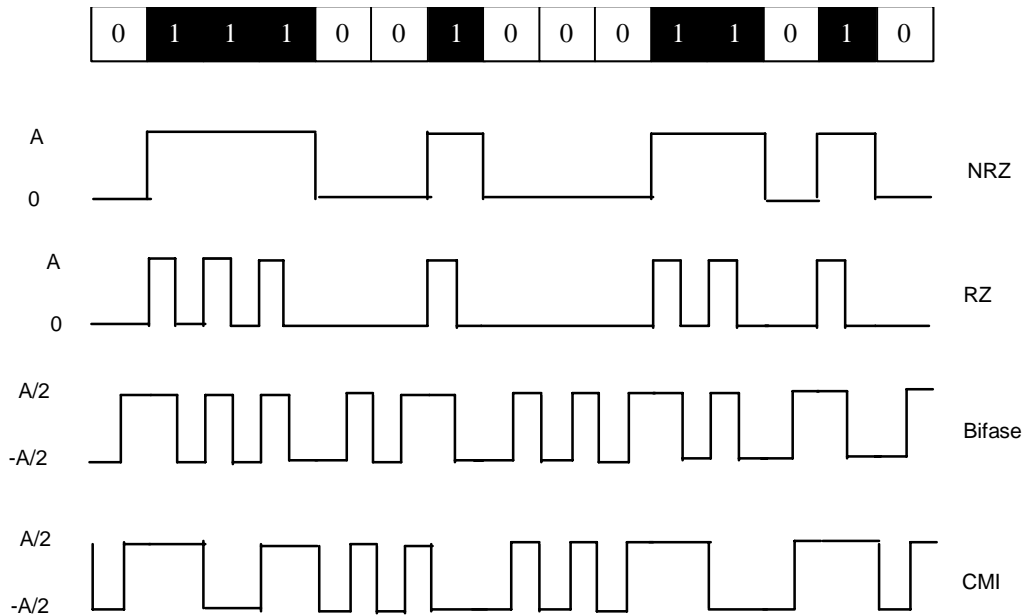



Figura 3.20 Diferentes tipos de códigos binários.

Embora tenha a vantagem de ser relativamente simples o código NRZ não satisfaz nenhuma das propriedades referidas anteriormente, não sendo, por conseguinte, apropriado para ser usado como código de linha em transmissão digital. O código RZ apresenta um conteúdo de temporização melhor do que o NRZ, mas necessita de uma banda de transmissão mais elevada.

Os códigos bifase (Manchester) e CMI (*coded mark inversion*) conseguem garantir um espectro com componentes nulas na frequência zero, assim como um conteúdo de temporização elevado, à custa de um aumento muito significativo da largura de banda do sinal. Como consequência é contraproducente a sua utilização sobre meios de transmissão metálicos, onde como se sabe, a largura de banda disponível é muito limitada, sendo no entanto, apropriados para a transmissão em fibra óptica. No código bifase tantos os 0s, como os 1s são codificados com uma transição a meio do período de bit ($T_b/2$). O código CMI é um código binário que codifica as sequências de 1s usando pulsos com duração igual ao período de bit e com polaridades alternadamente positiva e negativa (regra da

bipolaridade), enquanto o zero é codificado com um pulso de duração $T_b/2$ com polaridade negativa, seguido de um outro com a mesma duração mas de polaridade positiva. O código CMI é especificado como código de linha para a 4ª hierarquia da hierarquia digital plesiocrona europeia, que como se irá ver no Capítulo 6 corresponde ao débito de 139.264 Mb/s.

Os códigos AMI (*alternate mark inversion*) e o código HDB3 (*high density bipolar*) são dois códigos ternários, que têm a vantagem de garantirem as propriedades referidas no início desta secção e requerem a mesma largura de banda de transmissão que o código NRZ. A codificação das sequências de 1s no código AMI obedecem à regra da bipolaridade referida anteriormente e os 0s são codificados com um nível zero. Note-se ainda que os pulsos associados aos 1s podem ocupar todo o intervalo de bit (100 % *duty cycle*), ou só metade desse intervalo (50 % *duty cycle*).

Uma desvantagem do código AMI é a ausência de conteúdo de temporização em longas sequências de zeros. Este inconveniente é colmatado recorrendo ao código HDB3, onde sequências de quatro zeros são codificadas com 000V, ou com B00V. A primeira palavra consiste em três intervalos de bit sem pulso (000), seguidos por um pulso, que representa uma violação bipolar (ou seja, uma violação da regra da bipolaridade que impõe a seguir ao símbolo -1, o símbolo +1, e vice versa). A segunda palavra consiste num pulso que mantém a alternância bipolar (B), seguida de dois intervalos de bit sem pulso (00) e terminando com um pulso com violação bipolar (V). A decisão de substituir uma sequência de quatro zeros, por 000V ou B00V é feita de modo a garantir que o número de pulsos B entre violações é ímpar. O algoritmo de substituição está sintetizado na tabela 3.2. O código HDB3 pertence à família dos códigos BNZS (*bipolar* N-*zero subsitution*), onde sequências de *N* zeros são substituídas por sequências de bits especiais, contendo pelo menos uma violação da regra de bipolaridade, e é usado na codificação das hierarquias plesiocronas europeias de 2.048 Mb/s, 8.448 Mb/s e 34.368 Mb/s.

Tabela 3.4 Regra de formação do código HDB3

polaridade do pulso anterior	Número de pulsos desde a última subsituição	
	Impar	**Par**
-	**000-**	**+00+**
+	**000+**	**-00-**

Tabela 3.5 Regra de formação do código 2B3B

Bits de entrada	Modo 1	Modo 2
00	000	111
01	010	101
10	100	001
11	110	001

Os códigos de blocos tipo *m*B*n*B, também são muito usados em aplicações na Ethernet, particularmente na Ethernet de 100 Mb/s, 1 Gb/s e 10 Gb/s. Estes códigos convertem blocos de *m* bits em palavras de código de *n* bits, que são escolhidas de modo a garantir as propriedades referenciadas para os códigos de linha. Na tabela 3.5, descrevem-se, como exemplo, as regras de formação do código 2B3B. Como se pode ver, para cada grupo de

bits de entrada têm-se duas palavras de código (modos), que são usadas alternadamente, de modo a garantir que a componente de corrente contínua é nula.

3.4 Problemas

3.1 Considere um cabo de pares simétricos caracterizado pelos parâmetros primários dados na tabela 3.1. Para a frequência de 10 kHz determine o erro cometido quando calcula o coeficiente de atenuação com base na eq. (3.8), em vez de usar a expressão exacta.

3.2 Demonstre que, no caso em que a reactância indutiva da linha é muito maior que a resistência o coeficiente de atenuação pode ser aproximada pela equação (3.11).

3.3 Um cabo de pares simétricos é caracterizado pelas seguintes constantes primárias: L=0.55 μH/m, C=50 pF/m, R=0.02Ω/m e G=10 nS/m. Para esse cabo determine:
a) A impedância característica da linha.
b) O coeficiente de atenuação expresso em Np/km e dB/km.
c) A velocidade de propagação.
d) Admitindo que a linha está adaptada e que se aplica no seu início um sinal de amplitude de 1 V, qual é a amplitude do sinal ao fim de 100 km?

3.4 Numa rede local usam-se pares simétricos de 0.9 mm (19 AWG). Os seus parâmetros distribuídos são dados por: R=55 Ω/km, L=0.62 mH/ km e C=52 nF/km. Admita que dois assinantes envolvidos numa chamada estão ambos distanciados da central local de 4 km. Sabendo que as perdas da central são de 0.5 dB, e que aos terminais do microfone do telefone emissor é gerado um sinal com uma tensão de 100 mV à frequência de 1 kHz determine:
a) A atenuação no lacete de assinante.
b) A tensão na terminação analógica da central.
c) A tensão aos terminais do auscultador do telefone receptor.

3.5 Deduza a equação que permite calcular a largura de banda de uma fibra óptica (Eq. 3.23).

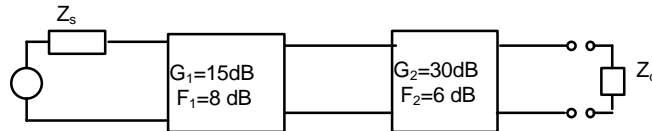
3.6 Considere uma ligação usando fibra óptica monomodal na distância de 100 km operando a 1.2 Gbit/s. O comprimento de onda de trabalho é de 1.55 μm, a atenuação da fibra é de 0.25 dB/km (incluindo as juntas) e o parâmetro de dispersão é igual a 16 ps/nm/km. A sensibilidade do receptor é de -30 dBm e a margem de segurança é de 4 dB. Nesta situação determine:
a) A largura espectral máxima do laser a usar.
b) A potência óptica de emissão.

3.7 Admite que entre Porto e Faro (distância de cerca de 570 km) se tem um sistema de transmissão em fibra óptica operando a 155 Mbit/s. Tendo presente que o índice de

refracção da fibra óptica é de 1.5, determine o número de bits que estão em trânsito entre essas duas cidades.

3.8 Determine o valor da potência de ruído térmico num canal de fonia (4 kHz) e num canal de vídeo (6 MHz) admitindo que a temperatura ambiente é de 290 K, exprimindo o resultado em dBW e dBmp (no primeiro caso).

3.9 Dois amplificadores ligados em cascata têm as seguintes especificações: T_1= 3T_o, G_1=10 dB, F_2=13.2 dB e G_2=10 dB. Se a sua largura de banda for de 100 kHz e a temperatura equivalente de ruído na entrada do primeiro igual a T_1= 10T_o, determine a potência do sinal na entrada, para garantir uma relação sinal-ruído de 30 dB: Exprima essa potência em pW, dBW e dBm (T_o = 290 ºk).

3.10 Considere a seguinte cadeia de quadripolos:

Admitindo que Δf=10 kHz e Z_s=600 Ω, determine:
a) A potência disponível de ruído à saída e a potência transmitida a uma carga de Z_c=75Ω.
b) A relação sinal/ruído na carga Z_c=75 Ω, admitindo que o valor eficaz do sinal sinusoidal é de 2 mV.

3.11 Considere um filtro de recepção passa-baixo ideal, cujo módulo da função de transferência é dado por:

$$|H(f)| = \begin{cases} 1 & se \ |f/B| \le 1 \\ 0 & se \ |f/B| > 1 \end{cases}$$

Suponha que tem à entrada do filtro uma fonte de ruído térmico (branco e Gaussiano) com uma densidade espectral de potência de $\mathcal{N}_0(f)=\mathcal{N}_0/2$. Sabendo que $\mathcal{N}_0$ =-110 dBm/Hz e B=600 kHz determine:
a) A densidade espectral de potência de ruído à saída do filtro.
b) A potência média de ruído à saída do filtro.
c) A largura de banda equivalente de ruído do filtro.

3.12 Considere agora a família de filtros passa-baixo realizáveis, com uma função de transferência em módulo dada por:

$$|H(f)| = \frac{1}{\sqrt{1+(f/B)^{2n}}}, \qquad n = 1,2,3,...$$

também designados por filtros passa-baixo de Butterworth de ordem n.
a) Represente esquematicamente o módulo dessa função em dB.
b) Repita o problema anterior para a família de filtros dada.

3.13 Admita que o coeficiente de atenuação de um cabo coaxial tem uma variação com a frequência aproximada pela equação 3.13. Nesta situação, determine e represente graficamente:

a) A função de transferência do igualador que compensa a atenuação e a distorção da linha;

b) A função de transferência do igualador, que conduz a uma interferência intersimbólica nula (factor de excesso de banda 0), tendo como código de linha o código NRZ.

3.14 Considere um sistema de transmissão digital a 34 Mbit/s por cabo coaxial na distância de 150 km. A atenuação do cabo é de 10 dB/km e usam-se 30 repetidores idênticos. Determine a potência a injectar na entrada do cabo para garantir na recepção uma taxa de erros binários de 10^{-7}, considerando que o filtro receptor é do tipo coseno elevado com $\alpha=0$ e se usa um código NRZ polar. Admita que:

a) Os repetidores são amplificadores que compensam exactamente a atenuação do cabo e apresentam um factor de ruído de 8 dB.

b) Os repetidores são regeneradores.

3.15 Considere um sistema de transmissão digital operando a 8 Mb/s, na distância de 75 km, com 25 repetidores. O meio de transmissão é o cabo coaxial de calibre 1.2/4.4 m, caracterizado pelos parâmetros a=0.07, b=5.15 e c=0.005. O código de linha usado é o AMI e o filtro receptor apresenta uma resposta Butterworth de 2ª ordem.

a) Determine a atenuação total do cabo;

b) Atendendo aos resultados da alínea anterior, determine a potência a injectar na entrada do sistema de modo a garantir uma taxa de erros de 10^{-6}, admitindo que os repetidores são amplificadores com um ganho de 32 dB e um factor de ruído de 7 dB;

c) Admitindo que cada repetidor é um regenerador, determine a potência a injectar na entrada do cabo, tendo presente que cada regenerador compensa uma atenuação de 40 dB, e o seu sistema de amplificação apresenta um factor de ruído de 7 dB.

d) Desenhe um regenerador apropriado para esse sistema.

3.16 Um sistema de transmissão digital operando a 1.544 Mbit/s usa como meio de transmissão um cabo de pares simétricos, cuja atenuação é de 10 dB/km a 1 MHz e como código de linha o código bipolar. Este sistema opera com uma taxa de erros binários de 10^{-6} e é limitado pela diafonia. Quais as modificações que teria de realizar para obter uma taxa de erros de 10^{-8}?

3.5 Referências

[Fo83] P.G. Fontolliet, *Systémes de Télécommunications*, Dunod, 1983.

[Ga03] M. Gagnaire, *Broadband Local Loops for High-Speed Internt Access,* Artech House, Inc., 2003.

[Hi79] M.T. Hills, B. G. Evans, *Telecommunications System Design*, GeorgeAllen&Unin Ltd. 1973.

[Re79] W. D. Reeve, *Subscriber Loop Signaling and Transmission Handbook*, IEEE Press 1995.

[Si04] D. R. Smith, *Digital Transmission Systems*, Kluwer Academic Publishers, 2004.

4 Aspectos da infra-estrutura da rede fixa de acesso

4.1 Estrutura da rede fixa de acesso

4.1.1 Rede convencional

Uma parte significativa da rede fixa local ou de acesso é constituída pela infra-estrutura de cobre que liga a instalação do assinante à central local, designada por **lacete local** (ou lacete de assinante). Essa infra-estrutura é constituída quase exclusivamente por pares simétricos, cujas características de transmissão foram objecto de análise no capítulo anterior.

Dentro da solução tradicional saem do **repartidor principal**[37], da central telefónica local vários cabos de pares simétricos, cada um constituído por milhares de pares simétricos, que vão ser separados em feixes para um determinado número de áreas de serviço (ver Figura 4.1). Estas áreas de serviço podem ter diferentes dimensões, desde algumas dezenas de quilómetros quadrados nas áreas urbanas, até algumas centenas nas áreas rurais. O número de pares por área de serviço é planeado com antecedência de modo a ultrapassar o número de assinantes na respectiva área, permitindo servir futuros utilizadores num curto espaço de tempo.

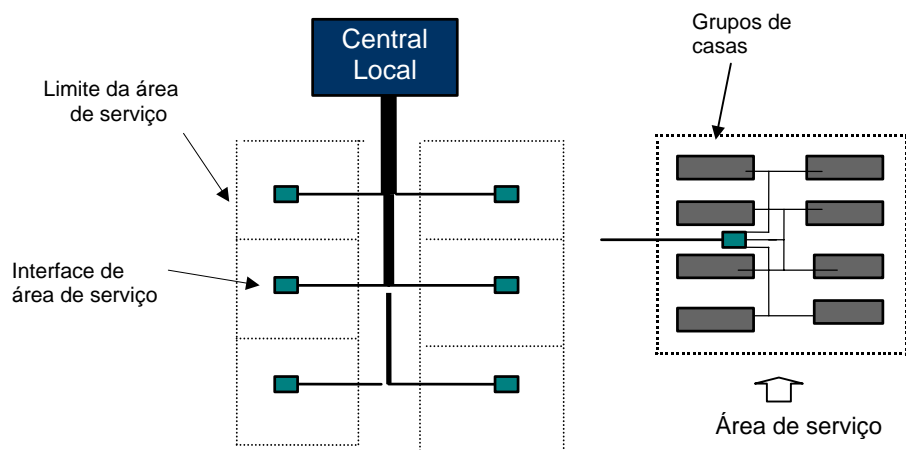



Figura 4.1 *Rede de distribuição local.*

Os cabos telefónicos que saem do repartidor principal da central de comutação local podem conter por vários milhares desses pares. Esses cabos, designados por **cabos de alimentação**, terminam normalmente num sub-repartidor, também designado por ponto de sub-repartição, de onde saem **cabos de distribuição** de menores dimensões, que por sua vez terminam nas caixas de distribuição de assinante, como se mostra na Figura 4.2.

[37] Também designado por MDF (*main distribution frame)*

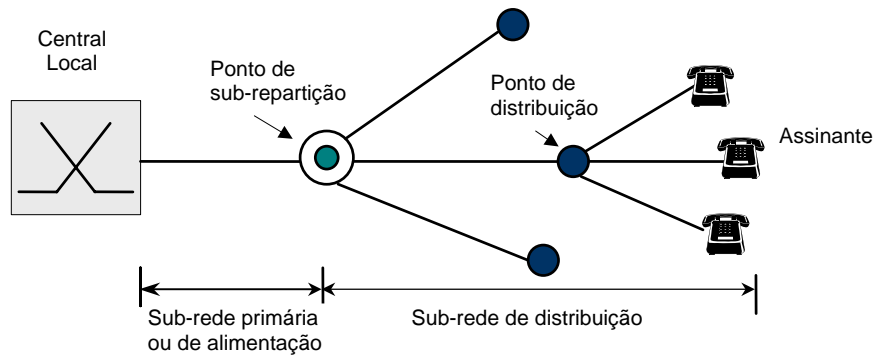



Figura 4.2 Arquitectura típica de uma rede de acesso

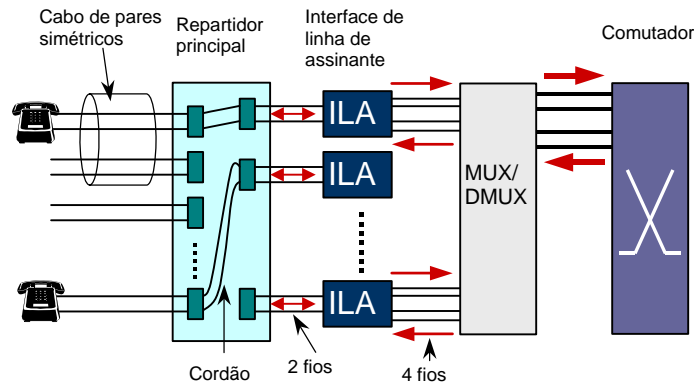



Figura 4.3 Estrutura simplificada da central local.

O repartidor funciona como terminação dos cabos de alimentação, e o mesmo tempo permite fazer a interligação entre os pares simétricos e as interfaces de linha de assinante (ILA) da central, usando para isso dispositivos designados por cordões (ver Figura 4.3). Assim, para desligar o telefone de um assinante basta remover o cordão que interliga o par simétrico atribuído a esse assinante. A ILA cujas funcionalidades irão ser detalhadas na Secção 4.2.2, é usada, nomeadamente, para passar de 2 para 4 fios (e vice-versa) e fazer a conversão analógica/digital e assim gerar o sinal PCM a 64 kbit/s correspondente ao sinal vocal gerado pelo equipamento terminal. Depois dessa operação, o sinal digital gerado é agregado, juntamente com outros sinais, num multiplexador, antes de ser encaminhado para a rede através do comutador. A estrutura e o funcionamento do multiplexador e do comutador irão ser estudados, respectivamente, nos Capítulos 5 e 6. No sentido inverso, o sinal digital na saída do comutador, vai ser desmultiplexado, antes de ser convertido do domínio digital para analógico e transmitido no par simétrico. A informação digital entre o multiplexador/desmultiplexador (também designados por MUX/DEMUX) é efectuada usando tramas E1 (2.048 Mb/s), ou T1 (1.544 Mb/s)[38].

[38] O E1 é usado na Europa e o T1 nos Estados Unidos

A atenuação máxima admissível no lacete local analógico é de 8 dB (@1 kHz). Assim, o projecto da rede local envolve uma escolha criteriosa do calibre do cabo de modo a garantir aquele objectivo. Normalmente, o calibre dos pares simétricos usados nos cabos de alimentação é menor do que nos restantes cabos. Assim, uma linha telefónica típica consiste, normalmente, em diferentes secções de cabo com diferentes calibres, ligados entre si, podendo ainda incluir derivações, que não são mais do que pares de cobre em circuito aberto, inseridos de modo a aumentar a flexibilidade da rede. A grande variação na composição das diferentes linhas telefónicas conduz a uma grande variabilidade das características de transmissão de ligação para ligação. Esta variabilidade é ainda acentuada pelo facto de haver partes da rede instaladas há várias dezenas de anos, enquanto outras partes são de instalação mais recente e por isso fazem uso de cabos de qualidade superior.

Ao longo das redes de acesso longas (maior do que 5 km), como aquelas que são usadas para servir zonas rurais, era habitual colocar bobinas de modo a compensar os efeitos capacitivos dos cabos (pupinização) e assim aplanar a resposta em frequência da linha na banda da voz. Só que, como já se referiu, a introdução dessas bobinas vai impedir a migração para serviços digitais, pelo que essa solução foi abandonada.

Outra alternativa possível para a rede de acesso consiste em colocar **unidades remotas** próximas de grupos concentrados de assinantes. Esta solução permite aliviar os cabos de alimentação e simultaneamente reduzir as perdas de transmissão das linhas telefónicas servidas pelas unidades remotas. Neste caso, essas unidades incluem as ILAs e o equipamento de multiplexagem/desmultiplexagem representado na Figura 4.3. Este equipamento multiplexa no tempo (TDM) vários canais telefónicos, e o sinal resultante é transportado para a central telefónica local usando por exemplo fibra óptica.

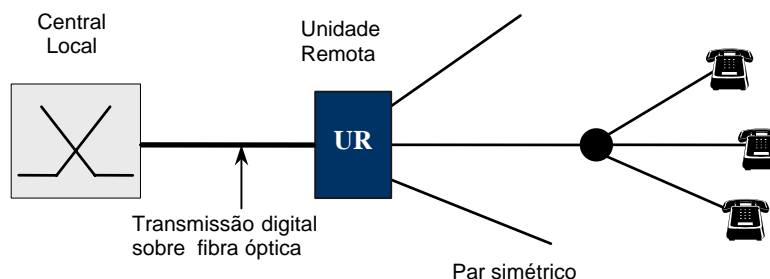



Figura 4.4 *Utilização de unidades remotas na rede de acesso.*

As unidades remotas podem funcionar no modo concentrado ou não concentrado. No primeiro caso a unidade remota, designa-se por **concentrador**[39]**,** e é ligada a um número de assinantes superior ao número de canais da via de transmissão digital que liga a UR à central local. Refira-se, como exemplo, o caso em que a via de transmissão suporta 30 canais e a UR serve 240 assinantes. Tal sistema apresenta um factor de concentração de 8. O modo concentrado é sobretudo interessante para áreas com tráfego reduzido, na medida em que a concentração envolve a existência de bloqueio (como se irá analisar no Capítulo 6), isto é, quando o número de pedidos de chamada é superior ao número de canais da via

[39] A diferença essencial entre concentradores e comutadores é que contrariamente aos segundos os primeiros não efectuam a comutação directamente entre dois terminais.

de transmissão, existem um certo número de pedidos que não são atendidos. Probabilidades de bloqueio entre 0.1 e 0.5% são perfeitamente aceitáveis. No modo não concentrado não há bloqueio, porque o número de canais disponíveis no sinal multiplexer é igual ao número de assinantes. Note-se ainda que, quando essa unidade funciona como concentrador, para além do equipamento já referido, é também necessária a presença de algum equipamento de comutação.

4.1.2 Tipos de acessos à Internet

O acesso à Internet pode-se classificar como directo ou indirecto. O acesso directo é usado, normalmente, pelos utilizadores empresariais e requer a presença de um *route*r IP nas instalações do utilizador. O acesso indirecto é usado pelos utilizadores domésticos e pressupõe o estabelecimento de uma ligação temporária ou permanente a um ISP (*Internet Service Provider*)[40], sendo a ligação feita por intermédio de um ponto de presença ou POP (*point of presence*). O acesso indirecto pode ser de banda estreita ou de banda larga. O acesso indirecto de banda estreita é realizado através de modems que operam na banda da voz, que permitem enviar o fluxo de informação gerado pelo computador através da linha telefónica. A escolha do POP apropriado para uma determina ligação é feita pela rede telefónica de um modo transparente para o utilizador, sendo posteriormente a chamada encaminhada para esse POP. O POP da rede telefónica pode por sua vez estar ligado aos POPs dos diferentes ISPs, usando circuitos alugados, ou canais virtuais permanentes estabelecidos usando uma rede ATM (constituída por vários comutadores), como se exemplifica na Figura 4.5.

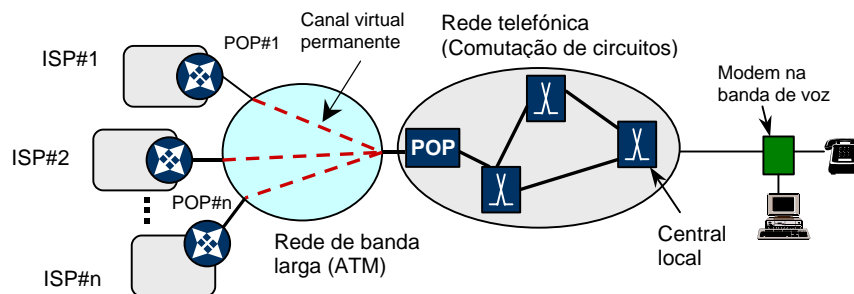



Figura 4.5 Acesso indirecto de banda estreita.

O acesso da banda larga pode ser feito usando a tecnologia ADSL, ou outras soluções alternativas que irão ser detalhadas ao longo deste capítulo. No caso em que se usa a tecnologia ADSL, a rede de acesso, que faz uso da infra-estrutura de pares simétricos existente entre os utilizadores e a central local, é interligada aos ISPs, usando também uma rede de banda larga baseada geralmente no ATM. A rede de acesso neste caso inclui para além dos modems ADSL nas instalações do cliente, os multiplexadores de acesso

[40] ISPs são entidades comerciais que proporcionam serviços aos utilizadores finais da Internet. Estes serviços abrangem a enorme variedade de servidores Web disponíveis em todo o mundo e além deste também servidores de áudio e de vídeo.

DSL ou DSLAM[41] localizados na mesma estação onde se encontra a central local. Cada DSLAM interliga várias centenas de modems ADSL à rede da banda larga. Para permitir a existência do serviço de banda estreita (seja analógico ou RDIS) no mesmo par simétrico, os sinais de banda estreita e banda larga são combinados na estação local usando um filtro apropriado, sendo a situação revertida nas instalações do assinante, de modo a separar os sinais correspondentes ao serviço telefónico dos sinais correspondentes aos serviços de banda larga, como se ilustra na Figura 4.6.

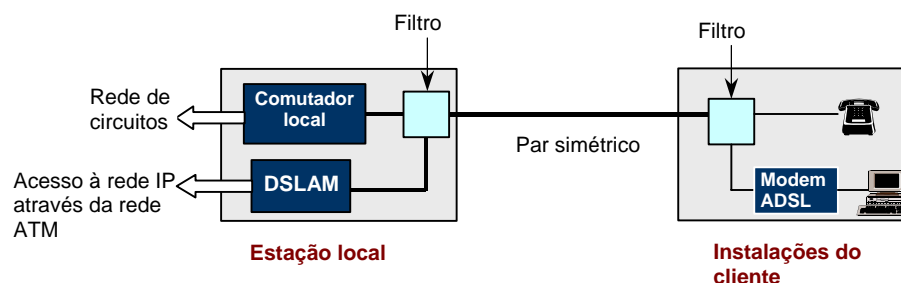



Figura 4.6 Acesso local com ADSL.

4.1.3 Desagregação do lacete local

O termo desagregação do lacete local descreve a obrigação[42] dos operadores da rede fixa históricos[43] de alugarem (ou venderem) a sua infra-estrutura de acesso baseada nos pares simétricos a operadores de telecomunicações alternativos. Este conceito foi desenvolvido nos Estados Unidos em meados da década de 90 do século passado e foi posteriormente exportado para a Europa, e tem por objectivo incentivar a concorrência no sector das telecomunicações. A desagregação pode ser parcial ou completa. Na desagregação completa o operador alternativo tem acesso pleno ao par simétrico do operador histórico, podendo por isso oferecer, quer serviços telefónicos, quer serviços de acesso à Internet de banda larga. Para isso, o operador alternativo tem de instalar na estação do operador histórico a sua central de comutação local e o seu DSLAM, como se exemplifica na Figura 4.7, os quais serão ligados aos pares simétricos dos seus clientes através de uma configuração apropriada do repartidor principal. Na desagregação parcial o operador da rede só entrega ao operador alternativo, ou o sinal banda base, ou o sinal de banda larga com a banda base filtrada e por isso a oferta dos serviços é parcial.

A política de desagregação do lacete local pode, no entanto, contribuir para dissuadir os operadores alternativos de investirem em novas tecnologias para rede local e certamente que também não motivará o operador histórico a fazer investimentos de vulto num segmento da rede, que por obrigação terá de partilhar com a concorrência. Esta situação pode conduzir à estagnação tecnológica do acesso e mesmo em algumas situações à degradação da infra-estrutura existente. Face a estes dados, os organismos reguladores terão certamente de reanalisar esta política de modo a criar as condições para a modernização tecnológica da infra-estrutura de acesso, cuja face mais visível será a

[41] DSL Access Multiplexer.
[42] Imposta pelos reguladores como é o caso da ANACOM em Portugal.
[43] Também designados por incumbentes.

introdução de fibra óptica, ainda com diferentes gradientes de penetração. O primeiro passo no sentido de alterar a política de desagregação foi dado de novo pelos Estados Unidos, com uma ordem emitida no último trimestre de 2004 pela FCC[44] que em traços gerais determina o seguinte [GR05]:

- Qualquer infra-estrutura de acesso baseada na FTTH[45] criada de raiz, não fica sujeita à obrigação da desagregação, ou seja não necessita de ser partilhada.
- Em qualquer infra-estrutura de acesso FTTH que resulte da substituição da infra-estrutura de cobre existente, somente a banda base destinada ao tráfego de voz terá de ser partilhada.
- Em todas as outras situações continuará a vigorar a politica de desagregação.

Os frutos desta alteração do panorama regulamentar nos Estados Unidos não deixaram de se fazer sentir quase imediatamente com grandes operadores como AT&T e Verizon a anunciarem investimento muito volumosos dirigidos para o acesso óptico, com resultados já visíveis, como seja por exemplo a interligação óptica de cerca de 3 milhões de casas no ano de 2005 levada a cabo só pelo último operador.

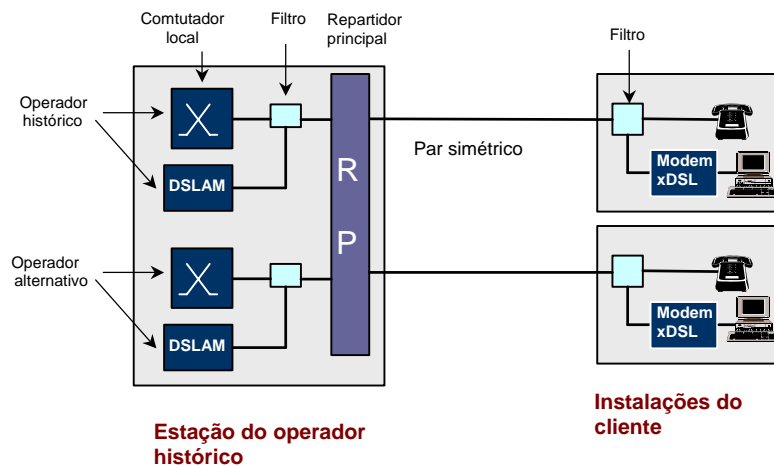



Figura 4.7 Desagregação completa do lacete local.

4.2 Equipamento terminal para lacete de assinante analógico

4.2.1 Equipamento terminal de assinante

O equipamento terminal de assinante analógico inclui, nomeadamente, os telefones residenciais, as cabines públicas, os terminais de telecópia, etc. Os modems usados para a transmissão de dados na banda de voz são também fontes de informação analógica, na medida em que estes dispositivos são usados para adaptar a informação digital às características de transmissão dos canais analógicos usados para a transmissão de voz.

---

[44] *Federal Communication Commission*
[45] *Fiber-to-the-home*

O equipamento terminal convencional do assinante é constituído pelo telefone. Cada telefone é alimentado por bateria central (situada na central de comutação local), que fornece uma tensão contínua de -48 V (valor típico).

Quando o telefone está no descanso este não é percorrido por nenhuma corrente significativa, pois a impedância do lacete local é muito elevada. Quando o telefone é levantado estabelece-se um fluxo de corrente I_S no lacete (ver Figura 4.8). O valor dessa corrente depende da tensão de alimentação e da resistência da linha do lacete.

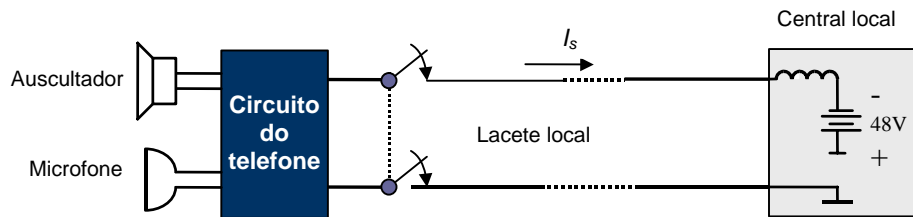



Figura 4.8 Telefone alimentado com bateria central.

A resistência do lacete inclui a resistência da linha telefónica e a resistência do próprio telefone, sendo a resistência da linha telefónica dependente do calibre dos pares simétricos e do seu comprimento. Valores típicos para a resistência máxima do lacete variam entre os 1250 e 1800 Ω, enquanto os valores típicos para a corrente de lacete se situam entre os 20 e os 100 mA.

No sentido de uniformizar a qualidade de serviço proporcionada pelo telefone será conveniente que as correntes vocais cheguem à central de comutação local sensivelmente com o mesmo nível, qualquer que seja o comprimento do lacete. Nesse sentido, os telefones actuais incluem um dispositivo regulador (ou compensador), que permite garantir em certa medida, que a corrente que percorre o microfone é independente do comprimento do lacete de assinante. Esses reguladores são, normalmente, varistores (resistências não lineares), cuja resistência decresce à medida que a corrente no lacete aumenta, ou seja, que o seu comprimento diminui. Em síntese, o efeito do varistor é variar a resistência do telefone em torno do seu valor nominal situado entre os 100 e 200 Ω.

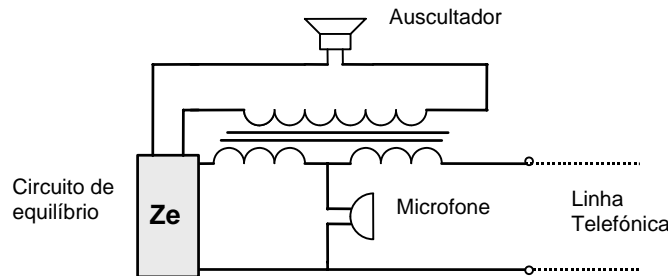



Figura 4.9 Conversão de dois para quatro fios usando um híbrido.

Como mostra a Figura 4.8 no telefone é necessário fazer uma conversão de quatro fios (auscultador+microfone) para dois fios, pois no lacete de assinante os sinais correspondentes à emissão e recepção viajam no mesmo par simétrico. Essa conversão é feita usando um dispositivo denominado híbrido, que se representa de modo simplificado

na Figura 4.9. No caso em que há equilíbrio perfeito, ou seja, quando a impedância do circuito de equilíbrio (Z_e)[46] é idêntica (em módulo e fase) à impedância apresentada pela linha (Z_l), a corrente (vocal) gerada pelo microfone é repartida em duas partes iguais, que fluem pelo circuito primário do híbrido com sinais contrários, fazendo com que a corrente gerada no secundário onde está ligado o auscultador seja nula. Deste modo, este circuito também se costuma designar por circuito anti-efeito local, já que os sinais vocais gerados pelo microfone não afectam o auscultador.

Os telefones modernos usam no circuito de equilíbrio circuitos apropriados para compensar a impedância de linha, para diferentes comprimentos desta e, assim, garantir um equilíbrio perfeito. Contudo, os telefones são, normalmente, projectados com um certo desequilíbrio, de modo a que o utilizador possa ouvir a sua voz e, assim, ter a percepção de que o telefone está activo. Na Figura 4.10 representa-se um esquema simplificado de um telefone. Nesse esquema podem-se identificar os seguintes elementos: a **campainha**, o **gancho** e respectivos contactos, o **marcador** (disco ou teclado), **compensador** (varistor), **emissor** (microfone), **receptor** (auscultador) e o **híbrido** com o respectivo **circuito de equilíbrio.**

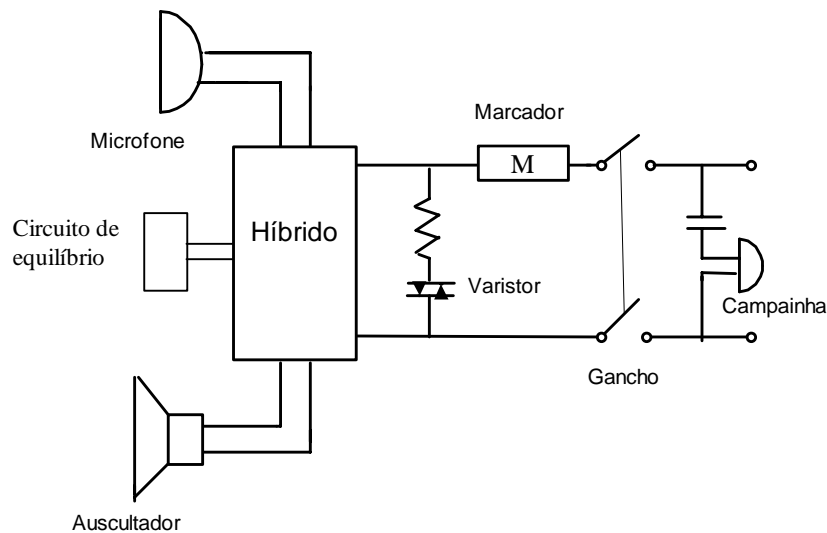



Figura 4.10 Estrutura simplificada de um telefone.

A campainha é activada aplicando uma tensão alterna de 75 Vrms à frequência de 25 Hz. A capacidade em série com a campainha é usada para evitar que a corrente contínua que percorre o telefone, quando os contactos associados ao gancho estão fechados, atravesse a campainha. O marcador é responsável pela sinalização de lacete. No caso dos telefones de disco, este marcador não é mais do que um interruptor, que interrompe a corrente contínua um número de vezes idêntico ao dígito marcado (sinalização decádica). No caso da sinalização multifrequência a marcação de um número activa dois geradores de frequências diferentes de acordo com a matriz representada na Figura 4.11. Por exemplo, marcação do número 2 envolve a geração dos tons de 697 Hz e 1336 Hz. Este tipo de sinalização tem a vantagem relativamente à decádica de requerer um tempo de

[46] A norma ESTI define que a impedância de equilíbrio é constituída por uma resistência de 270 Ω ligada a uma resistência de 750 Ω em paralelo com uma capacidade de 150 nF.

atraso de marcação menor e uniforme para todos os números, mas em contrapartida requer um receptor apropriado na central local.

No caso do RDIS ou da voz sobre IP os telefones são digitais, ou seja a conversão A/D e D/A ocorre no próprio equipamento do telefone. Tal implica que os tipos de sinalização referidos não podem ser usados nessas redes, exigindo-se em contrapartida a aplicação de protocolos de sinalização adequados.

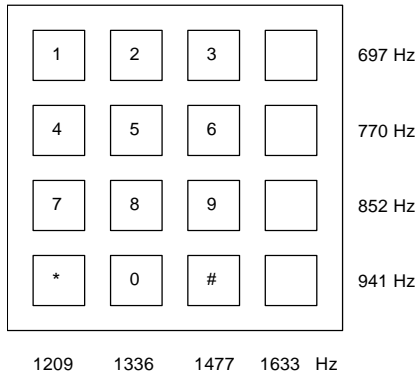



Figura 4.11 Tons usados na sinalização multifrequência.

4.2.2 Equipamento terminal da central local

Uma central de comutação local digital que serve lacetes de assinante analógicos é responsável pela realização de um conjunto de funções que podem ser sintetizadas pelo acrónimo *BORSCHT,* ou seja, alimentação (*Batery*), protecção contra sobre-tensões (*Over-voltage protection*), geração da corrente de toque da campainha do telefone (*Ringing*), supervisão do estado da linha de assinante (*Supervision*), conversão dos sinais analógicos para digitais e vice-versa (*Coding*), conversão de 2 para 4 fios (*Hybrid*) e vários tipos de teste (*Testing*). Essas funções são, normalmente, implementadas num circuito designado por **interface de linha de assinante** (ILA), que se representa na Figura. 4.12.

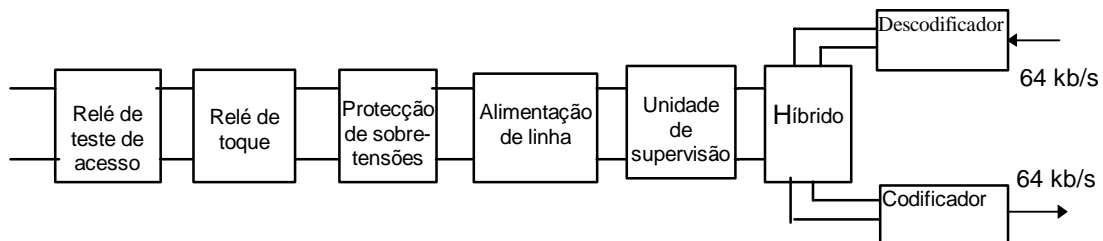



Figura 4.12 Interface de linha de assinante analógica numa central de comutação digital.

As diferentes funções realizadas pela ILA são descritas sinteticamente a seguir:

- **Alimentação de linha**: A alimentação do telefone com uma tensão de -48 V é realizada recorrendo a uma bateria central. Nas centrais analógicas a potência era

fornecida aos telefones dos dois assinantes envolvidos numa chamada através da utilização de pontes de transmissão. Na Figura 4.13 apresenta-se o circuito de uma dessas pontes, o qual é conhecido por ponte de Stones.

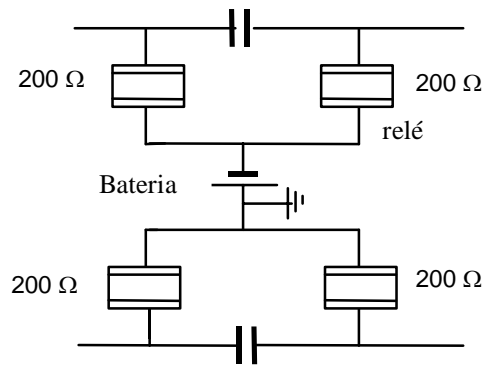



Figura 4.13 Ponte de transmissão de Stones.

Nesse circuito usam-se relés (bobinas enroladas em núcleo metálicos), para bloquear a transmissão dos sinais vocais para a bateria, já que estes dispositivos apresentam baixa resistência à corrente contínua e uma alta impedância às frequências vocais. A ponte de transmissão é completada pela utilização de condensadores em sério com os lacetes de assinante, de modo a isolar a nível de corrente contínua as duas linhas.

Contrariamente às centrais analógicas, as centrais digitais deixam passar unicamente as correntes e tensões correspondentes aos níveis lógicos digitais (Ex: 5 V). Assim, as funções correspondentes à ponte de transmissão são realizadas pela ILA, como se mostra na Figura 4.14 onde cada ILA realiza metade das funções correspondentes à ponte de transmissão.

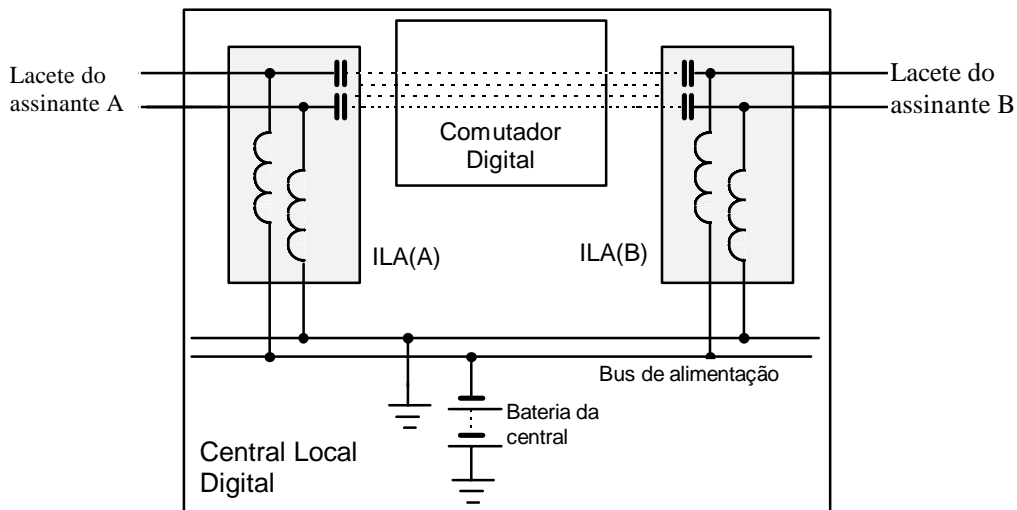



Figura 4.14 Mecanismo de alimentação numa central digital.

- **Protecção contra sobre-tensões:** Protecção do equipamento e do pessoal contra altas tensões originadas, quer por descargas atmosféricas, quer por cruzamento com linhas de alta tensão.

- **Toque da campainha**: A central local necessita de enviar um sinal de alerta para o telefone do assinante chamado, avisando-o do facto de estar uma chamada em espera. A frequência do sinal de chamada é de cerca de 25 Hz e a sua tensão *rms* é de 75 V. Este sinal está ligado durante 2 s e está desligado durante 4 s, e é obtido pela interrupção de um gerador de corrente partilhado por vários telefones.

- **Supervisão**: Como as centrais digitais não permitem um caminho metálico entre os assinantes envolvidos numa chamada, a supervisão do lacete de assinante é realizada na periferia da central, ou seja, na parte analógica do ILA, contrariamente às centrais analógicas, onde era realizada no interior da própria central. A função essencial da supervisão consiste em analisar o estado do lacete de assinante, detectando a presença ou a ausência do fluxo de corrente contínua nesse lacete, e converter este estado num sinal apropriado para ser interpretado pelo sistema de controlo da central. Esta tarefa requer um sensor (normalmente um relé) com capacidade para discriminar eficientemente, qualquer que seja o comprimento da linha telefónica, entre a corrente correspondente ao estado fora-do-gancho (telefone levantado) e a corrente resultante do ruído e das correntes de fuga.

- **Codificação**: Corresponde às funções A/D e D/A, as quais foram analisadas com detalhe no Capítulo 2.

- **Híbrido**: O híbrido é responsável pela conversão de 2 fios para 4 fios e vice-versa. O seu funcionamento já foi analisado a propósito do estudo do telefone. A única diferença a salientar é que o híbrido da ILA deverá ser projectado para um equilíbrio tão perfeito quanto possível.

- **Teste**: Esta função requer o acesso ao lacete local e aos circuitos da central para detectar possíveis falhas e proporcionar manutenção. Os testes devem ser feitos automaticamente, em horas de fraca utilização e com periodicamente.

4.3 Circuitos de 2 e 4 fios

Como já se referiu na rede local usa-se transmissão a dois fios. Contudo, para distâncias de transmissão superiores a cerca de 50 km é necessário separar fisicamente as duas direcções de transmissão. Há duas razões para isso: primeiro, os circuitos longos requerem amplificação e regeneração e tanto os amplificadores como os regeneradores são dispositivos unidireccionais; segundo, por razões de economia, as chamadas telefónicas de longa-distância são multiplexadas usando multipexegem por divisão no tempo (TDM)[47], e essa multiplexagem requer que os sinais nas duas direcções sejam

[47] Usando as hierarquias plesiocronas, ou síncronas que irão ser estudadas no Capítulo 5.

enviados em *time-slots* distintos. A comutação digital também é feita usando circuitos com quatro fios, implicando, como se viu, uma conversão de 2/4 fios na interface de linha de assinante. Essa conversão é feita usando um híbrido. Um parâmetro importante usado para caracterizar um híbrido é a atenuação transhíbrido, A_t. Esta atenuação pode-se decompor em duas parcelas (ver Figura 4.15):

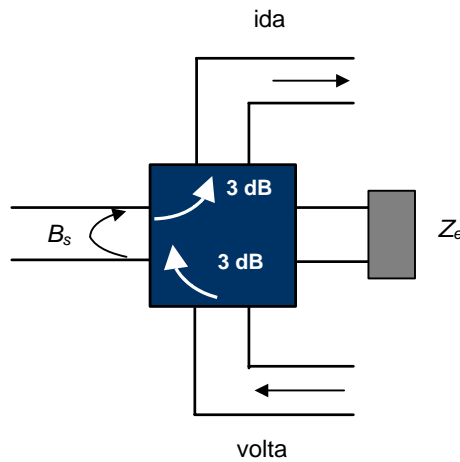
- dois termos de 3 dB devidos à natureza própria do híbrido e correspondentes à passagem do terminal a 2 fios de volta (do terminal a 4 fios) para o terminal a 2 fios e deste para o terminal a 2 fios de ida (do terminal a 4 fios);

- a atenuação de equilíbrio, B_s, que resulta da desadaptação de impedâncias da linha (Z) e do circuito de equilibrio (Z_e).

Como consequência dessa desadaptação, uma parte da potência recebida do terminal a 4 fios é reflectida, sendo o coeficiente de reflexão dado por

$$\Gamma = \frac{Z - Z_e}{Z + Z_e} \tag{4.1}$$

A atenuação de equilíbrio, obtém-se directamente a partir do coeficiente de reflexão, vindo

$$B_s = 20\log_{10}\left|\frac{1}{\Gamma}\right| \tag{4.2}$$

Figurar 4.15 Definição das parcelas contribuintes para a atenuação transhíbrido.

A atenuação transhíbrido vale, portanto, idealmente A_t=6+B_s dB, ou na prática A_t=7+B_s dB, sendo a diferença devida às perdas internas do híbrido. Note-se que, para Z=Z_e vem B_s=∞ dB, mas que para a linha em aberto (Z=∞), ou em curto-circuito (Z=0), vem B_s=0 dB.

Considere-se em seguida um sistema de transmissão de longa distância com circuitos de 4 fios. Como se representa na Figura 4.16, nesse sistema usam-se amplificadores para

compensar a atenuação do híbrido e do meio de transmissão. A atenuação total entre os pontos a dois fios é dada por

$$A_2 = 6 - G \text{ dB} \quad (4.3)$$

onde G é o ganho líquido de um dos lados do circuito a quatro fios (ganho total dos amplificadores menos as perdas do meio de transmissão). Outro parâmetro importante é a atenuação do anel, que é dada por

$$A_a = 2(A_2 + B_s) \text{ dB} \quad (4.4)$$

A não idealidade dos híbridos faz com que o circuito representado na Figura 4.16 se comporte como uma malha de realimentação, a qual se torna instável se:

- a atenuação do anel apresentar um valor negativo (em dB), ou seja, se a malha apresentar ganho;
- se o sinal reflectido depois de uma volta completa na malha estiver em fase com o sinal principal, ou seja quando a desfazagem total for múltipla de 360º.

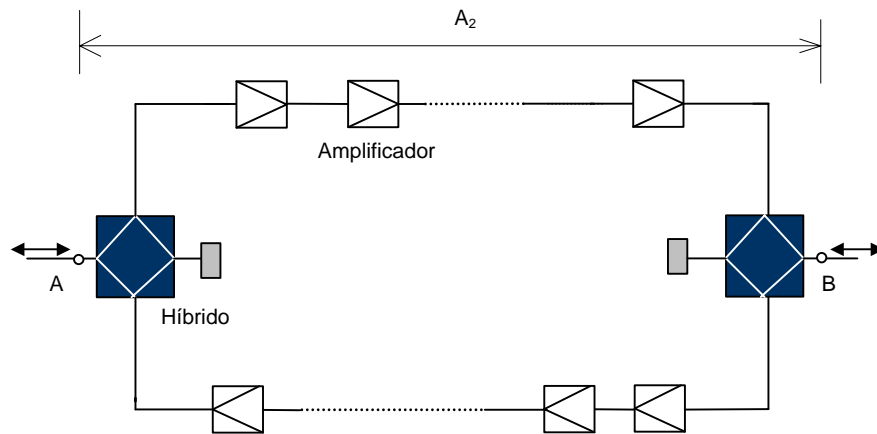



Figura 4.16 Sistema de transmissão a 4 fios.

Como consequência, podem surgir oscilações autosustentadas para as frequências do sinal em que se verificam essas condições. A amplitude dessas oscilações é unicamente limitada pela saturação dos amplificadores. Essa saturação irá inevitavelmente conduzir à distorção dos sinais transmitidos nesse sistema. Este fenómeno de auto excitação (*singing*) é, assim, extremamente perigoso e deve ser evitado a todo o custo. Para isso, é necessário garantir que a atenuação do anel é positiva. Para garantir uma margem de segurança (designada, normalmente, por margem de estabilidade) toma-se em geral

$$A_a \geq 6 \text{ dB} \quad (4.5)$$

Nestas condições tem-se

$$A_2 \geq 3 - B_s \text{ dB} \tag{4.6}$$

No caso limite, em que a linha está em curto-circuito, ou em circuito aberto (B_s=0 dB), ter-se-á

$$A_2 \geq 3\,\text{dB} \tag{4.7}$$

Repare-se que, a atenuação do anel é uma grandeza aleatória, que flutua em torno do seu valor nominal. As causas dessas flutuações são variadas, podendo referir-se como exemplo, a variação do ganho do equipamento com a temperatura e com o envelhecimento, desadaptações, variação do ganho dos circuitos entre a frequência de teste e as outras frequências vocais, etc. Essas flutuações são, normalmente, modeladas usando uma distribuição normal, cuja variância aumenta com o número de circuitos ligados na cadeia. Deste modo, o valor de A_2, deve ser seleccionado, de modo a que, mesmo nas piores condições de funcionamento ainda se consiga garantir uma margem de estabilidade conveniente. Uma regra simples, adoptada por alguns operadores é a seguinte:

$$A_2 = 4.0 + 0.5n \tag{4.8}$$

onde n representa o número de troncas (circuitos de 4 fios) presentes na ligação.

4.4 Supressores e canceladores de eco

4.4.1 Origem do eco

Uma outra consequência nefasta do facto de a atenuação transhíbrido não ser infinita é o aparecimento de sinais parasitas designados por ecos. Existem dois mecanismos distintos associados à geração dos ecos, os quais estão esquematizados na Figura 4.17. O eco de falante, que corresponde ao locutor ouvir uma versão atrasada da sua própria voz, enquanto no eco de ouvinte, o ouvinte ouve uma versão atrasada da voz do locutor.

A atenuação do eco de falante (desprezando a atenuação do circuito a dois fios) é dada por

$$A_{ef} = 2A_2 + B_s \tag{4.9}$$

enquanto o atraso é dado por

$$\delta T_{ef} = 2(T_2 + T_4) \tag{4.10}$$

onde T_2 e T_4 são, respectivamente, o tempo de propagação no circuito a 2 fios e no circuito a 4 fios. Por sua vez, para o eco do ouvinte tem-se uma atenuação igual a (desprezando de novo a atenuação a 2 fios)

$$A_{eo} = 2(A_2 + B_s) \tag{4.11}$$

enquanto o atraso é dado por

$$\delta T_{eo} = 2T_4 \tag{4.12}$$

O efeito subjectivo do eco depende criticamente do seu atraso e atenuação. Para pequenos atrasos, o eco de falante constitui uma perturbação desprezável desde que a atenuação do eco seja razoável (≥ 6 dB), sendo, neste caso, praticamente indistinguível do efeito local do telefone. Para atrasos de 40 ms ou superiores, o eco do falante constitui uma perturbação muito séria, a não ser que seja fortemente atenuado.

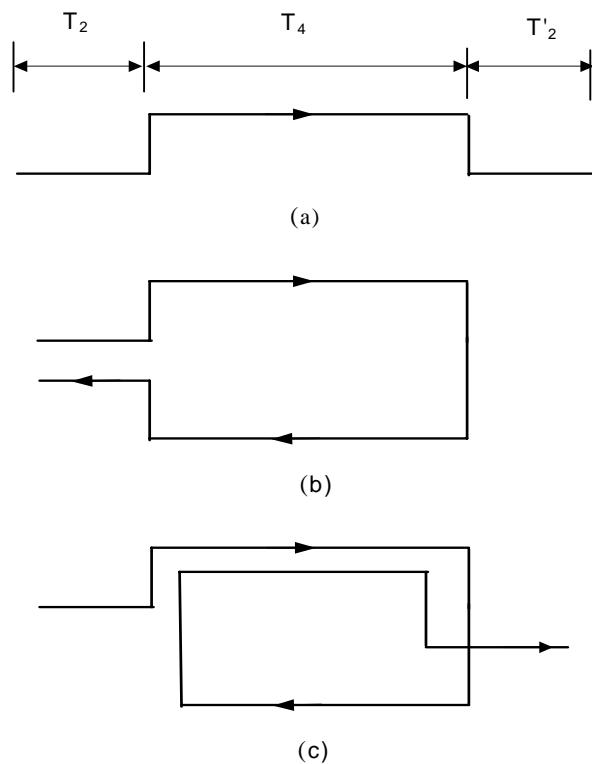



Figura 4.17 Tipos de ecos numa rede telefónica. (a) Ligação telefónica simplificada; (b) Percurso do eco de falante; (c) Percurso do eco de ouvinte.

Na Tabela 4.1 apresentam-se os valores da atenuação para o percurso do eco considerado satisfatório para um ouvinte médio, e, como se pode ver, quando o atraso de propagação aumenta, também tem de aumentar a atenuação, para produzir o mesmo efeito subjectivo.

Como se pode ver pelas equações anteriores, o eco de ouvinte é mais atenuado do que o eco de falante, sendo por isso, menos problemático do que o último, particularmente na transmissão de voz. No caso da transmissão de dados usando modems na banda da voz o efeito do eco de ouvinte é mais problemático e por isso deverá ser tido em consideração nos estudos de desempenho destes serviços [Co98].

Tabela 4.1 Atenuação no percurso do eco de falante (em dB) considerada satisfatória pela média dos ouvintes, em função do tempo de propagação num sentido.

Tempo de propagação (ms)	Atenuação (dB)
10	11.1
20	17.7
30	22.7
40	27.2
50	30.9

O atraso introduzido na transmissão da informação, devido não só ao tempo de propagação, mas também ao processamento de sinal e às funções de encaminhamento, vai influenciar para além do eco também o nível de interactividade entre os interlocutores envolvidos numa ligação. Um atraso inferior a 150 ms por sentido de transmissão permite um bom nível de interactividade sem afectar significativamente a conversação, enquanto atrasos entre os 150 e 300 ms conseguem garantir um nível de interactividade aceitável. Atrasos superiores a 400 ms por sentido de transmissão devem ser evitados, porque com estes atrasos a conversação se torna praticamente *half-duplex.*

O problema do atraso é um dos condicionalismos mais relevantes na transmissão de voz sobre IP. Os processos de amostragem, codificação, segmentação e formação dos pacotes podem originar atrasos até cerca de 64 ms. Se se juntar outros 64 ms para a recepção dá um total de 128 ms, o que deixa simplesmente 22 ms, para a transmissão sobre a rede, se a exigência for os 150 ms [Gr04]. Notar que na rede para além do atraso de propagação, há o atraso associado à comutação de pacotes, o que faz com que o requisito de 150 ms não seja fácil de atingir na transmissão de voz sobre IP.

4.4.2 Controlo do eco

Uma conclusão que se tira imediatamente das equações (4.9) e (4.11) é que os efeitos do eco podem ser reduzidos se se aumentar a atenuação do percurso A_2. Como a atenuação do sinal é proporcional a A_2, enquanto a do eco é proporcional a $2A_2$ (note-se que A_2 é expresso em dB), tem-se, assim, uma melhoria de relação sinal-eco. Contudo, para circuitos de comprimento elevados (superior a cerca de 3000 km) esta técnica não pode ser usada, porque origina um nível de sinal muito baixo no receptor. Para estes circuitos o controlo dos ecos pode ser feito usando dois dispositivos chamados **supressores de eco** e **canceladores de eco.**

Os supressores de eco são circuitos automáticos que efectivamente restringem a bidireccionalidade das comunicações telefónicas para garantirem uma atenuação suficientemente elevada no percurso do eco. A ideia base consiste em interromper, ou

atenuar significativamente um sentido de transmissão, logo que um detector de nível de voz constata a presença de um sinal de voz no outro sentido de transmissão.

A presença do supressor de ecos efectiva um regime de transmissão *half-duplex,* inviabilizando, por conseguinte, uma transmissão *full-duplex.* Existem duas situações em que a transmissão *half-duplex* é problemática. Em primeiro lugar, quando os dois interlocutores associados a uma ligação telefónica falam simultaneamente. Neste caso, a presença do supressor de ecos impede, na realidade, a comunicação bidireccional, interrompendo, conforme as circunstâncias, ora um, ora outro dos interlocutores.

A segunda situação problemática ocorre na transmissão de dados. Como, geralmente, os sistemas de transmissão de dados requerem um canal de retorno para pedidos de retransmissão em caso de detecção de erros, seria necessário desactivar esses supressores para uma transmissão efectiva.

Mesmo em face desses inconvenientes, os supressores de eco foram usados com êxito durante muitas décadas, principalmente nas ligações terrestres em que o tempo de atraso de ida e volta não excede os 100 ms.

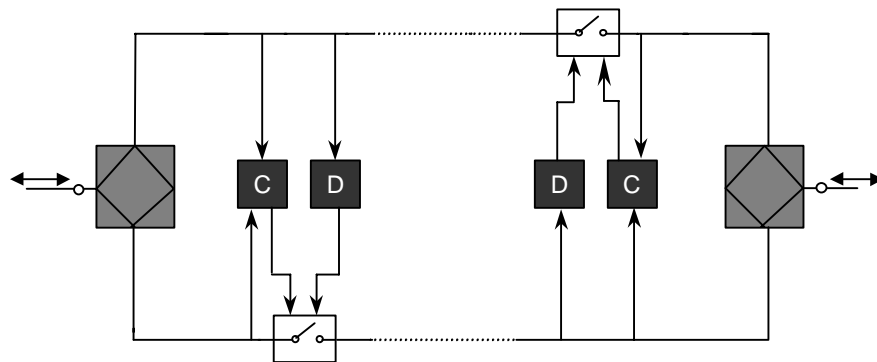



Figura 4.17 Princípio de funcionamento de um supressor de ecos (D: detector de nível de voz, C: comparador de nível de voz nos dois sentidos de transmissão).

Com o desenvolvimento das comunicações via satélite esses tempos de atraso podem atingir várias centenas de ms (~ 300 ms) e por isso foi necessário desenvolver uma nova tecnologia, capaz de enfrentar com sucesso o problema dos ecos, a tecnologia dos canceladores de eco. Neste caso, em vez de interromper (atenuar) o percurso do eco o cancelador sintetiza uma réplica do eco, que é subtraída ao sinal recebido. Se essa réplica for perfeita tem-se, assim, um cancelamento total do efeito do eco.

Para uma análise mais detalhada de um cancelador de ecos considere-se a Figura 4.17. O elemento chave desse dispositivo é um filtro adaptativo, que tem capacidade para sintetizar a função de transferência do percurso do eco, sendo ainda capaz de acompanhar as alterações resultantes das variações desse percurso, que ocorrem durante uma ligação telefónica. O sinal proveniente do interlocutor distante designa-se por $y(t)$, o do eco por $r(t)$ e o do interlocutor próximo por $x(t)$. O eco vai perturbar o sinal $x(t)$, como se mostra na Figura 4.17. O sinal $y(t)$ funciona como sinal de referência para o cancelador de ecos, sendo a partir deste que o filtro adaptativo sintetiza uma réplica do eco, designada por $\hat{r}(t)$. Esta réplica é subtraída do sinal $z(t)=x(t)+r(t)$, originando o seguinte sinal de erro:

$$\varepsilon(t) = z(t) - \hat{r}(t) = x(t) + r(t) - \hat{r}(t) \qquad (4.13)$$

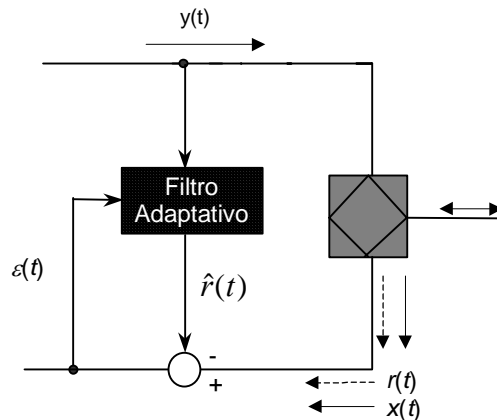



Figura 4.17 Princípio do cancelador de eco.

Repare-se que, no caso ideal em que a réplica é perfeita ($\hat{r}(t) = r(t)$) o efeito do eco é completamente suprimido, de modo que o sinal de erro é igual a $x(t)$. Nos casos práticos, tem de se ir ajustando dinamicamente a função de transferência do filtro adaptativo usando um algoritmo de adaptação apropriado, para sintetizar uma réplica tão perfeita quanto possível. Uma das possíveis estratégias para esse ajuste, consiste em minimizar o valor quadrático médio do sinal de erro $\varepsilon(t)$.

Nas redes telefónicas, normalmente, só se recorre ao uso de canceladores de ecos nas redes com atrasos mais significativos e estes são colocados logo a seguir ao híbrido na componente a 4 fios. Nas redes celulares ou de voz sobre IP o uso de cancelador de eco é quase obrigatório devido aos atrasos elevados, e esses canceladores são incluídos nas MSCs das redes celulares, ou nos *gateways* das redes VoIP que fazem a ligação às redes telefónicas.

4.4.3 Técnicas de duplexagem

Para conseguir transmitir os sinais digitais sobre uma infra-estrutura de cabos de pares simétricos é necessário recorrer a técnicas especiais para garantir a bidireccionalidade numa transmissão. As técnicas de transmissão bidireccional, ou de duplexagem usadas numa ligação a 2 fios são as seguintes:

- Duplexagem por divisão na frequência (DDF);
- Duplexagem por divisão no tempo ou TCM (*Time Compressed Multiplexing*);
- Cancelamento de eco (CE).

4.4.3.1 Duplexagem por divisão na frequência

A técnica de duplexagem por divisão no domínio da frequência (DDF) é uma técnica estabelecida já há muito tempo e consiste em dividir a largura de banda do canal, em duas bandas, uma para o percurso ascendente (utilizador-central local) e outra para o percurso descendente (central local-utilizador). Para evitar interferências entre as duas bandas será conveniente usar uma banda de guarda. Será de salientar que em termos de eficiência será conveniente escolher uma banda de guarda tão reduzida quanto possível, mas por sua vez a redução da interferência aconselha a aumentar essa banda, fazendo com que para projectar de modo adequado os filtros a usar seja necessário encontrar um ponto de equilíbrio.

4.4.3.2 Duplexagem por divisão no tempo ou TCM

A ideia subjacente a esta técnica consiste em separar as direcções de transmissão no domínio do tempo. A sua implementação é relativamente simples, baseando-se na transmissão alternada de blocos de informação num sentido e no outro sobre um único par simétrico, como se indica no diagrama de blocos da Figura 4.18.

No TCM a sequência binária gerada pela fonte de informação é fragmentada em blocos de informação (rajadas) constituídos por N bits. A duração de cada bloco é de $\Delta t = N / D_b$, onde D_b é o débito da sequência binária. Cada bloco é transmitido na linha ao ritmo de D_{b0}.

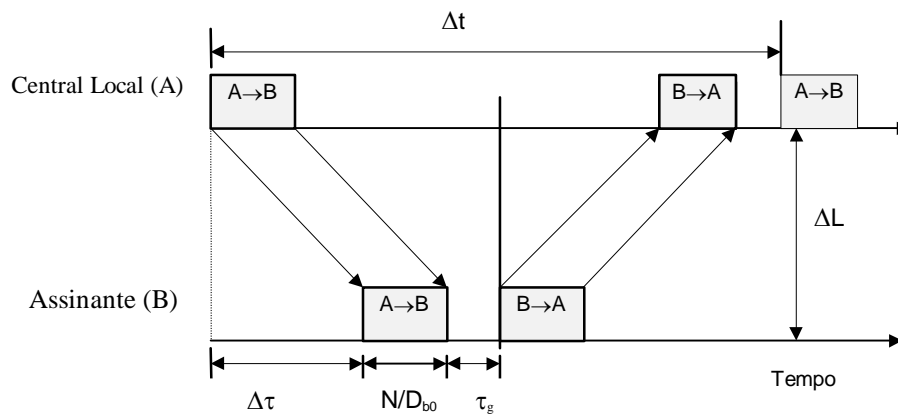



Figura 4.18 Transmissão de informação usando o TCM.

Seja $\Delta\tau$ o tempo de propagação na linha ($\Delta\tau = \Delta L / v_g$) e τ_g o tempo de guarda. Durante o intervalo de tempo Δt é necessário garantir uma comunicação bidireccional, ou seja, o débito binário na linha deverá ser tal que

$$\frac{N}{D_b} = 2\frac{N}{D_{b_0}} + 2\Delta\tau + 2\tau_g \tag{4.14}$$

o que permite ainda escrever

$$D_{b0} = \frac{2D_b}{1 - \frac{2}{\Delta t}(\Delta\tau + \tau_g)} \tag{4.15}$$

Como se conclui facilmente da equação anterior, o débito binário da linha terá de ser superior ao dobro do débito binário com que a informação é gerada, o que é sem dúvida uma limitação significativa desta técnica. Em contrapartida, se houver sincronização de todos os sistemas na central esta técnica é insensível à paradiafonia. Contudo, com a desagregação do lacete local não vai ser possível realizar a sincronização de todos os sistemas e deste modo esta vantagem do TCM poderá perder-se.

4.4.3.3 Tecnologia de cancelamento de eco

A tecnologia de cancelamento de eco (CE) permite fazer a transmissão bidireccional simultânea a dois fios, usando um cancelador de ecos do tipo estudo na secção 4.4.2, de modo a eliminar as perturbações originadas pela não idealidade do híbrido ou por reflexões ocorridas na linha.

4.5 Lacete digital de assinante

4.5.1 Limitações da tecnologia usada

Como já se viu a rede de acesso ou local corresponde à parte da rede situada entre o assinante e a central local. A infra-estrutura responsável pela transmissão nessa rede é baseada fundamentalmente em tecnologia de cobre, ou seja cabos de pares simétricos. Como se viu no Capítulo 3, as características eléctricas desses cabos são dependentes da frequência e das dimensões dos condutores. Para a transmissão digital as principais características de interesse são: atenuação, velocidade de grupo, impedância característica, diafonia e ruído impulsivo e gaussiano. A diafonia, ou seja a interferências eléctricas num par simétrico induzida pelos pares vizinhos é um dos factores mais limitativos da transmissão digital bidireccional nesses cabos. Através da fórmula de Shannon pode-se determinar a limitação na capacidade devida à diafonia. Essa fórmula pode escrever-se na seguinte forma:

$$C = B_c \log_2(1 + S/N) \quad \text{bit/s} \tag{4.16}$$

onde C representa a capacidade do canal, B_c a largura de banda do canal e S/N a relação sinal-ruído devida unicamente à diafonia (a potência da diafonia é calculada usando a equação (3.56). Com base nessa equação, representa-se na Figura 4.19 a capacidade de transmissão de um par simétrico com o diâmetro de 0.5 mm em função do seu comprimento. Daí pode-se concluir que só a diafonia limita o comprimento a valores de 3 km para débitos binários da ordem dos 750 kb/s.

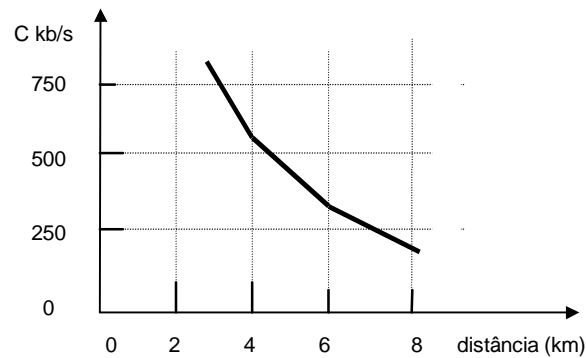



Figura 4.19 Capacidade de um par simétrico com um diâmetro de 0.5 mm em função do seu comprimento entrando unicamente em conta com a diafonia.

A transmissão digital sobre a rede de acesso com infra-estrutura convencional é usada fundamentalmente em três aplicações:

- Transmissão dados usando modems na banda de voz;
- Rede Digital com Integração de Serviços;
- Lacete Digital de Assinante Assimétrico ou ADSL (*Asymmetric Digital Subscriber Line*).

4.5.2 Transmissão de dados usando modems na banda da voz

A transmissão de dados usando modems foi o primeiro serviço digital a ser proporcionado usando a infra-estrutura de cabos de pares simétricos. Como os circuitos telefónicos convencionais eram analógicos e os computadores originam uma informação digital era necessária uma interface entre os dois. Essa interface designa-se por *modem,* que é uma contracção das palavras *modulator* e *demodulator.*

Um *modem* desempenha assim duas funções fundamentais. A sequência de bits originada por um computador é usada pelo *modem* para modular uma portadora apropriada para a transmissão sobre os cabos de pares simétricos. O *modem* da parte receptora desmodula essa portadora e regenera a sequência original. Notar que a transmissão dessa portadora modulada é feita no canal telefónico analógico, ou seja, entre os 300 e os 3400 Hz, daí o dispositivo em questão designar-se por modem na banda da voz.

As versões dos modems mais antigas como sejam os modems V.21, V.22 e V.22bis usavam como técnicas de modulação o FSK e o PSK e como técnica de duplexagem a divisão no domínio da frequência. Os modems seguintes desde o V.32 até ao V.34bis recorreram à modulação de amplitude em quadratura ou QAM e para garantir a transmissão bidireccional sobre a linha usavam cancelamento de eco (CE). Os modems referidos são modems analógicos. Por isso a informação por eles enviada sofre uma conversão A/D na central total e no destino o sinal digital terá de ser convertida de novo para o domínio analógico, para que seja possível realizar a desmodulação e recuperar a sequência de dados. O modem V.90, neste aspecto, apresenta-se como uma tecnologia

disruptiva, já que a transmissão no percurso descendente é feita sempre no domínio digital. Deste modo eliminam-se as conversões A/D e D/A o que permite atingir débitos de 56 kb/s. No percurso ascendente o funcionamento deste modem é idêntico ao do V.34 bis.

Tabela 4.1 Sucessivas gerações de modems de banda de voz

Geração	Data	Débito Binário (kbit/s)	Modulação+ Duplexagem
V.21	1978	0.3	FSK-DDF
V.22	1980	1.2	PSK+DDF
V22bis	1981	2.4	PSK+DDF
V.32	1985	9.6	QAM+CE
V.33	1988	14.4	QAM+CE
V.34	1990	28.8	QAM+CE
V.34bis	1995	33.6	QAM+CE
V.90	1998	33.6 (A); 56 (D)	QAM(A)+PCM(D)

Os modos de operação dos modems podem ser: *simplex, half-duplex* e *full-duplex.* No modo *simplex* os dados são transmitidos ou recebidos numa única direcção. No modo *half-duplex* a transmissão e a recepção de informação são realizadas alternadamente. Antes de se iniciar a transmissão numa direcção, o transmissor da direcção oposta deve ser desligado. No modo *full-duplex* o modem pode transmitir e receber dados simultaneamente. Nos circuitos a 2 fios este funcionamento é conseguido usando uma técnica de duplexagem apropriada, que no caso da DDF implica atribuir dois canais separados na frequência (banda inferior e banda superior) dentro da banda do canal telefónico, como se exemplifica na Figura 4.20. Um modem transmite na banda inferior e recebe na banda superior. O modem do extremo oposto transmite na banda superior e recebe na inferior. Nas linhas alugadas usa-se normalmente uma transmissão a quatro fios, sendo a bidireccionalidade garantida por circuitos fisicamente separados.

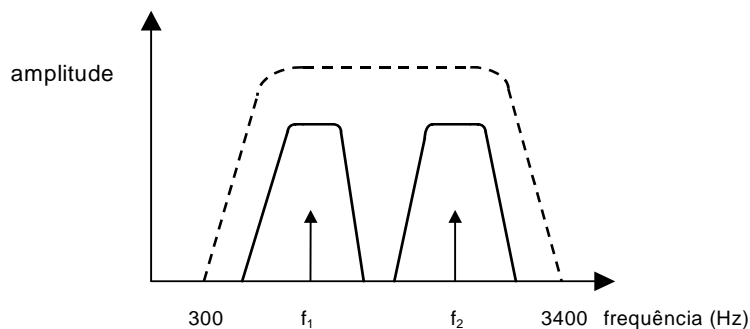



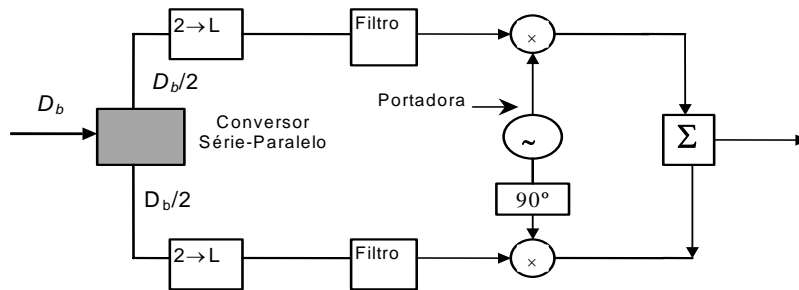
Figura 4.20 Plano de frequências dentro do canal telefónico usando DDF, sendo f_1 e f_2 as frequências das portadoras usadas, respectivamente, nos sub-canais de banda inferior e banda superior.

Para débitos binários superiores a 4800 bit/s torna-se difícil utilizar DDF em ligações *full-duplex*. Nessa situação recorre-se, normalmente, à técnica do cancelamento de eco. Esta técnica, que irá ser analisada quando se estudar a linha digital de assinante, permite que ambos os modems transmitam simultaneamente na mesma frequência e sobre uma ligação a dois fios.

Exemplo 4.1

O modem V.22 do ITU-T permite realizar um transmissão full-duplex a 1200 bit/s. É usada modulação 4-PSK e as duas portadoras têm respectivamente as frequências de 1200 e 2400 Hz. O débito de símbolo é de 600 baud, ou seja, as duas portadoras são moduladas a um ritmo de 600 baud. Como consequência a largura de **banda mínima** ocupada por cada uma das portadoras moduladas é de 600 Hz. Neste caso, a banda do subcanal inferior estender-se-ia de 900 a 1500 Hz e a superior de 2100 a 2700 Hz. Quais seriam os limites dos sub-canais se as portadoras moduladas ocupassem a largura de banda máxima (factor de excesso de banda igual a 1)? Qual é o débito binário máximo que é possível transmitir com este modem?

Exemplo 4.2

O M-QAM é uma técnica de modulação popular nos modems. Para calcular a largura de banda mínima necessária para transmitir um sinal valerá a pena recordar que esse sinal é obtido somando duas portadoras com a mesma frequência e em quadratura, cada uma delas modulada por um sinal multinível ($L=\sqrt{M}$). Basicamente um modulador de M-QAM tem a seguinte configuração:

A sequência binária proveniente da fonte com o débito binário de D_b é convertida em duas novas sequências com o débito de $D_b/2$. O conversor de 2 para L níveis transforma estas duas sequências em dois sinais PAM com L níveis e caracterizados pelo débito de símbolo de

$$D_s = \frac{D_b}{2\log_2 L} \text{ baud}$$

Estes sinais são em seguida filtrados por filtros passa-baixo. Assumindo que estes filtros têm uma largura de banda mínima (largura de banda de Nyquist com α=0), a largura de banda FI (depois de modulação) será

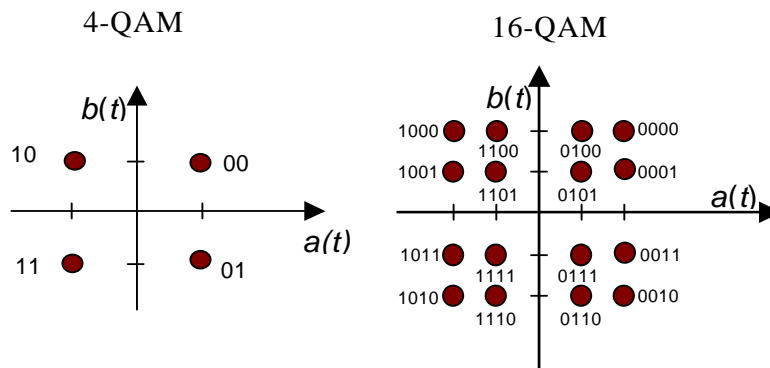
$$B_{FI}(\min) = D_s = \frac{D_b}{2\log_2 L} \text{ Hz}$$

A eficiência espectral do M-QAM pode-se assim calcular facilmente resultando:

$$\varepsilon = \frac{D_b}{B_{FI}(\min)} = 2\log_2 L \quad \text{bit/s/Hz}$$

Conclui-se assim que o 16-QAM tem uma eficiência espectral de 4 bit/s/Hz e o 64-QAM de 6 bit/s/Hz.

Exemplo 4.3
Um sinal QAM pode-se representar em termos de uma componente em fase e de uma componente em quadratura. Para uma portadora de frequência ω_i tem-se

$s(t) = \mathrm{Re}\{\sum_l c_l(t) e^{j\omega_i t}\} = \sum_l a_l(t)\cos\omega_i t - \sum_l b_l(t)\sin\omega_i t$, onde $c_l(t) = a_l(t) + jb_l(t)$. O conjunto dos valores de $c_l(t)$ e a sua representação no plano complexo designase por constelação do sinal M-QAM. As figuras abaixo representam as constelações do 4-QAM e do 16-QAM. Como se conclui facilmente à primeira modulação correspondem 2 bit/símbolo, enquanto a segunda é caracterizada por 4 bit/símbolo.

4.5.3 RDIS e digitalização do lacete de assinante

Uma das mais importantes inovações associadas à RDIS consistiu na deslocação da transmissão digital até ao assinante. Neste caso, o lacete do assinante designa-se por **lacete digital de assinante** ou **DSL** (*digital subscriber loop*), o qual serve de suporte para o acesso básico à RDIS. A Figura 4.21 mostra que este acesso requer a instalação de dois modems banda-base, um nas instalações do cliente designado por terminação de rede de tipo 1 (NT1[48]) e um outro na central local designado por terminação de linha (LT[49]). Para além disso, pode-se usar uma peça de equipamento adicional designada por terminação de rede de tipo 2 (NT2), que permite ao utilizador implementar, por exemplo, uma rede local (LAN) nas suas instalações. A ligação entre os elementos de rede NT1 e LT suporta um débito binário de 160 kb/s e é feita em banda base usando o código 2B1Q e a duplexagem da comunicação é assegurada, normalmente, usando canceladores de eco.

O acesso básico a 160 kb/s corresponde a um débito de 144 kb/s (dois canais B de 64 kb/s cada e um canal D de 16 kb/s) para o utilizador, mais um canal de 16 kb/s para funções de controlo de sincronização. Os canais B são usados para o tráfego de voz e dados e o canal D para sinalização. Como se referiu no Capítulo 1 está ainda normalizado um outro acesso designado por acesso primário, ou acesso 30B+D (30 canais B +1 canal D) no caso da Europa. Neste caso, a trama do sinal multiplexado é semelhante à do E1,

[48] *Network termination of type* 1
[49] *Line termination*

com o time-slot 0 destinado ao sincronismo e o time slot 16 ao canal D. Tem-se, assim, um débito binário de 2.048 Mb/s associado ao acesso primário. Neste caso, a transmissão é feita com quatro fios e o meio de transmissão pode ser a fibra óptica ou pares simétricos com repetidores para distâncias superiores a 1.5 km.

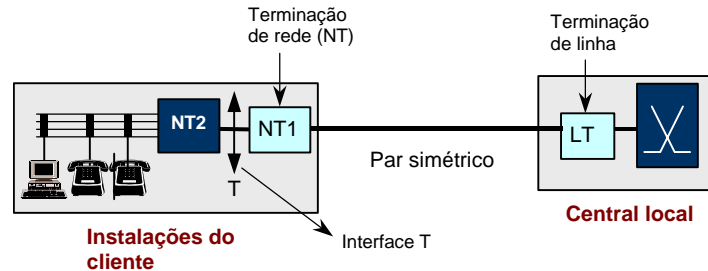



Figura 4.21 Configuração o acesso básico RDIS.

4.5.3.1 Tecnologias de transmissão para o DSL

Aquando da normalização do RDIS considerou-se que a transmissão no lacete de assinante deveria ser a dois fios de modo a aproveitar a infra-estrutura de cabos de pares simétricos da rede analógica já existente. Isto impõe a transmissão de débitos de 160 kb/s sobre os pares simétricos de cobre usados na transmissão telefónica analógica, o que não é uma tarefa fácil se se atender a que:

- A distância entre os assinantes e a central local é elevada podendo ir chegar a distâncias entre os 6 e 8 km;

- As atenuações são elevadas como consequência do elevado débito binário e das características dos cabos simétricos;

- A linha do assinante é em geral constituída por troços com características diferentes, nomeadamente, diferentes diâmetros, o que provoca reflexões devido à desadaptação de impedâncias;

- As características da linha são fortemente dependentes da frequência (por exemplo a impedância característica é cerca de 600 Ω para a frequências inferiores a 10 kHz e 150 Ω para as altas frequências.) Assim, uma linha adaptada para as baixas frequências poderá estar desadaptada para as altas-frequências, originando por conseguinte reflexões;

- Os cabos têm um elevado número de pares simétricos em muitos casos com mau isolamento, o que origina níveis de diafonia elevados e um aumento do valor da atenuação;

- Os cabos não foram dimensionados para transmissão a frequências tão elevadas, o que implica atenuações e distorções elevadas para essas frequências.

Para conseguir transmitir os sinais digitais correspondentes ao acesso básico da RDIS sobre uma infra-estrutura de cabos de pares simétricos é necessário recorrer a técnicas de duplexagem e a códigos de linha apropriados para a situação em causa. As técnicas de duplexagem usadas na DSL são o TCM e o cancelamento de eco. O TCM já foi estudado na secção 4.43 e baseia-se na separação das direcções de transmissão no domínio do tempo. Para realizar essa função é necessário usar no sistema um dispositivo de comutação que comuta sequencialmente do transmissor para o receptor, como se mostra na Figura 4.22.

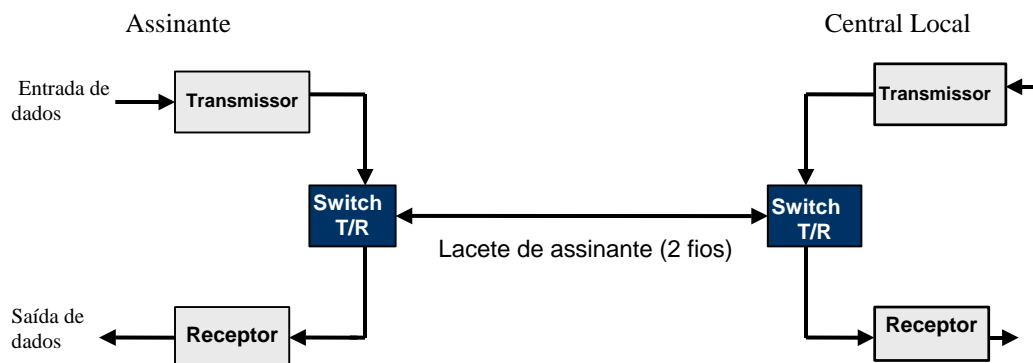



Figura 4.22 Transmissão bidireccional usando a técnica TCM.

Como se viu na secção 4.4.3.2 o débito binário da linha terá de ser superior ao dobro do débito binário com que a informação é gerada, fazendo com que no acesso básico do RDIS se requeiram débitos na linha da ordem dos 400 kbit/s, conduzindo a débitos a comprimentos máximos da linha da ordem dos 2 a 3 km.

Regras para o projecto de lacetes digitais de assinante usando cabos de pares simétricos:
Comprimento: ≤5.5 km para os cabos com diâmetro de 0.9, 0.6, e 0.5 mm e 4.5 km para os cabos com 0.4 mm
Atenuação: ≤ 42 dB à frequência de 40 kHz
Resistência do lacete ≤ 1300 Ω
Tipo de lacete: não carregado, ou seja não pupinizado.

4.5.3.2 Tecnologia de cancelamento de eco

A tecnologia de cancelamento de eco é baseada na transmissão bidireccional simultânea a dois fios com cancelamento automático do eco recebido, como se exemplifica na Figura 4.23. O sinal $x(t)$, recebido da linha, após a separação através do híbrido do sinal transmitido $y(t)$, é constituído por várias componentes sobrepostas ao sinal emitido pelo terminal remoto, sendo de destacar as seguintes:

- Uma pequena fracção do sinal transmitido $y(t)$ em virtude do híbrido não estar perfeitamente adaptado;

- Ecos múltiplos originados por reflexões em diferentes pontos do par simétrico, devido nomeadamente, a desadaptações de impedância, cujas causas já foram identificadas anteriormente;

- Ruído de origem térmica e ruído impulsivo originado por diferentes sistemas eléctricos, assim, como diafonia devido aos outros pares;

- Interferência intersimbólica resultante da transmissão num meio com banda muito limitada.

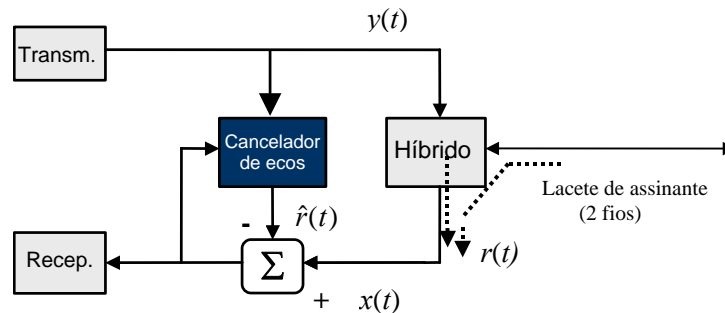



Figura 4.23 Estrutura do NT1 realçando a componente de cancelamento de eco.

Como o sinal recebido sofre uma forte atenuação, que pode atingir os 40 dB, pode acontecer que as perturbações acima referidas sejam da mesma ordem de grandeza deste, o que dificulta o processo de cancelamento de ecos e de igualação.

Como foi visto na secção 4.4.2 o cancelador de ecos é um filtro adaptativo cujo objectivo é simular a função de transferência do eco $r(t)$, de modo a originar uma réplica $\hat{r}(t)$ deste tão perfeita quanto possível, a qual vai ser subtraída ao sinal. No caso ideal, em que a réplica é perfeita, tem-se um cancelamento total do eco.

Para reduzir o efeito da interferência intersimbólica o receptor usa um igualador. Como as características de transmissão dos pares simétricos não são estáticas, variando constantemente devido a factores vários (temperatura, interferências electromagnéticas, etc.) esses igualadores terão de ser necessariamente adaptativos.

Esta tecnologia, ainda que mais complexa do que a anterior, é preferida actualmente, pois permite atingir comprimentos de ligação superiores, da ordem dos 6 a 7 km.

4.5.3.3 Códigos de linha

Normalmente o transmissor da Figura 4.23 inclui um codificador de linha e o receptor inclui o correspondente descodificador. Como a banda do canal é muito restrita, procura-se escolher códigos que conduzam a um débito de símbolo inferior ao débito binário. Códigos como o AMI e HDB3 não têm aqui tanto interesse como na transmissão digital de longa distância. Um código consideravelmente popular no DSL é o código 2B1Q, onde 2 bits são convertidos num símbolo quaternário, segundo a regra de codificação apresentada na Tabela 4.2.

Tabela 4.2 Código 2B1Q

Palavra Binária	Palavra quaternária
00	-3
01	-1
10	+3
11	+1

Para o débito de 160 kb/s tem-se um ritmo de símbolo de 80 kbaud, o que como se sabe conduz a um banda de canal mínima de 40 kHz.

4.5.4 ADSL

O Lacete Digital de Assinante Assimétrico ou ADSL (*Asymmetric Digital Subscriber Loop*) é uma tecnologia que proporciona transferência de informação de modo assimétrico, sendo o débito no sentido descendente (central-assinante) em regra superior ao débito no sentido ascendente (assinante-central).

O ADSL é uma tecnologia que usa tal como o DSL a infra-estrutura de cabos de pares simétricos, e surgiu no início dos anos 90 tendo por finalidade fornecer serviços de vídeo-a-pedido[50]. A aplicação desta tecnologia nesse contexto não teve contudo grande sucesso, porque requeria fortes investimentos em equipamentos para distribuição centralizada de vido digital e para além disso a qualidade do vídeo digital, obtido na altura usando a norma MPEG-1, era bastante fraca. O desenvolvimento da Internet e a necessidade de tecnologias que suportassem o acesso de banda larga levou a que por volta de 1997 o ADSL fosse redireccionado para aplicações de acesso à Internet.

O ADSL permite a transmissão simultânea no mesmo par simétrico de sinais digitais e do sinal associado ao serviço de telefonia analógico convencional. Os elementos de rede de um acesso ADSL estão representados na Figura 4.6. A transmissão da informação digital é feita num banda de frequências situada acima da banda tradicionalmente usada pelo serviço de telefonia. Note-se que para agregar/desagregar os dois sinais no mesmo par se requer a utilização de filtros repartidores/combinadores em ambas a extremidades do par. As funções de transmissão e recepção da tecnologia ADSL são realizadas por blocos (modems) designados de ATU[51]-C e ATU-R. O bloco ATU-C é responsável pelas funções de codificação e modulação (bem como descodificação e desmodulação) do sinal do lado da central local, enquanto o ATU-R se encarrega das mesmas funções do lado do cliente.

No sentido de combinar no mesmo par os sinais associados à transmissão no sentido ascendente, ou no sentido descendente a norma ADSL permite duas variantes. A primeira variante designada por DDF-ADSL usa duplexagem no domínio da frequência para separar as bandas correspondentes aos dois sentidos de transmissão (ver Figura 4.24). A comunicação ascendente tem lugar na banda de frequências situada entre os 25 e os 138 kHz, enquanto a comunicação descendente ocorre na banda de frequências situada entre os 150 e os 1104 kHz. Os débitos binários de transmissão alcançados dependem certamente das técnicas de modulação e das características da linha, tendo as normas

[50] *Vídeo-on-demand* segundo a notação anglo-saxónica.
[51] ADSL *transmission unit.*

apontado como alvo máximo um débito de 640 kb/s no sentido ascendente e de cerca de 6 MHz no sentido descendente. A segunda variante, designada por CE-ADSL, baseia-se no facto de a banda de frequências ascendente estar sobreposta à banda descendente, requerendo por isso a utilização de canceladores de eco, para separar os sinais nos dois sentidos. Como mostra a Figura 4.25 o tráfego descendente ocupa a banda entre os 25 e os 1104 kHz, que se sobrepõe com a banda ascendente localizada entre os 25 e 130 kHz. Esta segunda variante, embora mais cara do que a primeira, usa a banda de frequências disponível de modo mais eficiente e deste modo permite atingir débitos máximos de cerca de 800 kb/s para o tráfego ascendente e 8 Mb/s para o tráfego descendente. É claro, que o débito máximo alcançado vai depender também das características dos cabos, e neste caso particular, o efeito da paradiafonia na banda entre os 20 e 130 kHz pode ter uma acção muito limitativa.[52]

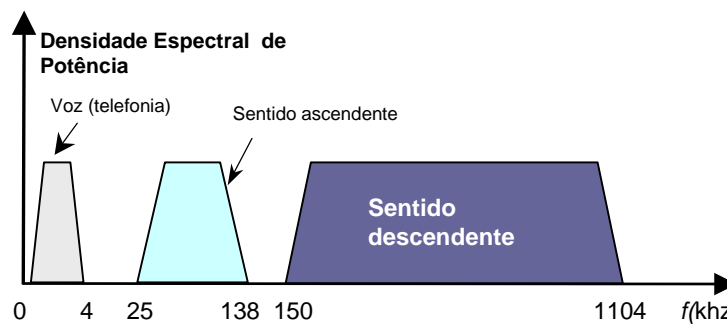



Figura 4.24 Ocupação da largura de banda de um sistema DDF-ADSL.

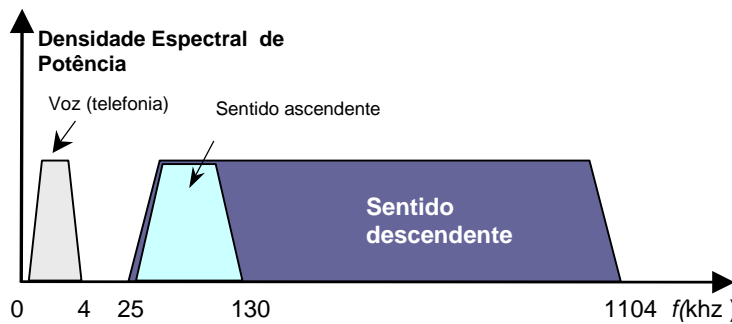



Figura 4.25 Ocupação da largura de banda de um sistema CE-ADSL.

Como já se referiu a diafonia pode ser um factor muito limitativo particularmente quando se usa a duplexagem baseada no cancelamento de eco. A atenuação é outro dos factores que condiciona grandemente o alcance dos sistemas ADSL. No capítulo 3 analisaram-se essas limitações e verificou-se que a atenuação dos pares simétricos cresce com a raiz da frequência. Esta variação faz, por exemplo, com que a um coeficiente da atenuação de 1dB/km à frequência de 1 kHz, corresponda um coeficiente da atenuação de 16 dB/km à frequência de 1 MHz. Estas limitações vão fazer com que o alcance dos sistemas ADSL se reduza à medida que o débito binário aumenta, como está patente na Tabela 4.3, que

[52] Note-se que o efeito da paradiafonia nos sistemas DDF-ADSL é desprezável, porque os dois sentidos de comunicação fazem uso de bandas de frequências diferentes.

inclui os valores das distâncias alcançáveis com um par simétrico de 0.5 mm de diâmetro para diferentes débitos do tráfego descendente. Note-se que o ADSL convencional só permite atingir os primeiros três valores dos débitos binários apresentados nessa tabela. Os dois débitos intermédios (9 e 13 Mb/s) são conseguidos usando a versão ADSL2, enquanto os últimos dois débitos só podem ser suportados com a tecnologia VDSL (*very high-speed digital subscriber line*).

Tabela 4.3 Alcance dos sistemas ADSL

Débito binário (Mb/s)	Distância (km)
1.5	5.4
2.0	4.8
6.0	3.6
9.0	2.7
13.0	1.35
26.0	0.9
52.0	0.3

A técnica de modulação mais usada nos sistemas ADSL é e **modulação multi-tom discreta**, conhecida usualmente pela designação de DMT (*discrete multitone)*. A ideia base desta modulação consiste em dividir a largura de banda operacional do ADSL num número elevado de sub-canais e usar em cada um dos sub-canais uma portadora (tom) diferente. A largura de banda associada a cada sub-canal deve ser suficientemente reduzida de modo a que o coeficiente de atenuação e o atraso de grupo possam ser considerados aproximadamente constantes nessa banda, fazendo com que o meio de transmissão possa ser considerado não distorcivo, como se viu na secção 3.21. Usualmente a banda de frequências do ADSL entre 0 e 1.104 MHz é dividida em 256 sub-canais (numerados de 1 a 256), o que conduz a uma largura de banda de 4.3125 kHz por sub-canal. Os primeiros 6 sub-canais correspondem à banda usada para suportar os serviços de telefonia analógicos e não são por isso usados. O tráfego ascendente ocupa 24 sub-canais, enquanto tráfego descendente ocupa 222 sub-canais na variante DDF-ADSL e 248 na variante CE-ADSL. As portadoras usadas no centro de cada sub-canal são por sua vez moduladas usando QAM.

A primeira fase do processo de modulação consiste em definir a dimensão da constelação QAM apropriada para cada sub-canal. Para calcular essa dimensão requer-se um processo de inicialização, através do qual o ATU-C e o ATU-R trocam sequências de teste de modo a avaliar a relação sinal-ruído de cada um dos 248 sub-canais (do 7 até ao 256) com interesse para o sistema ADSL. Com base no valor dessa relação sinal-ruído o modem ADSL adapta automaticamente a dimensão da constelação QAM a ser usada em cada sub-canal, em conformidade com o débito binário de transmissão requerido. A Figura 4.26 ilustra de modo esquemático o princípio descrito. A parte superior da figura descreve os valores médios da relação sinal-ruído medidos em cada um dos sub-canais, ao passo que a parte inferior apresenta um exemplo de uma possível atribuição de capacidade (bit/símbolo) pelos diferentes sub-canais. Como é de esperar, quanto maior for a relação sinal-ruído maior é a dimensão da constelação considerada. Os sub-canais nº 2 e 3 são os que apresentam a maior qualidade e por isso é-lhes atribuído o valor da

capacidade mais elevado, ou seja 5 bit/símbolo, o que corresponde ao 32-QAM. O sub-canal nº 5 como apresenta uma relação sinal-ruído muito baixa não é usado, enquanto os restantes canais são caracterizados por uma capacidade de 4 bit/s, requerendo, por isso, a utilização de 16-QAM.

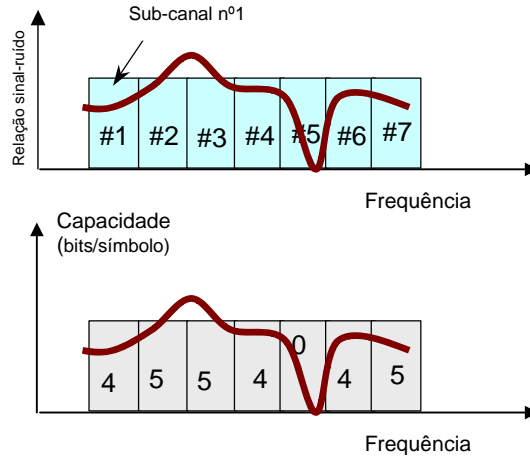



Figura 4.26 Princípio da atribuição de capacidade na modulação DMT.

A Figura 4.27 apresenta um esquema simplificado de um modulador DMT considerando a presença de *n* sub-portadoras (ou seja *n* sub-canais). A sequência digital com o débito binário D_b entra num conversor série-paralelo, sendo por este convertida em tramas constituídas por *M* bits. Os bits da trama são agrupados em *n* sub-blocos $\{B_1, B_2, ... B_n\}$, sendo a dimensão de cada sub-bloco determinada pela dimensão da constelação corresponde ao sub-canal associado a esse sub-bloco. Esses bits são ainda reordenados e por isso para além do conversor série-paralelo terá de existir uma memória-tampão adicional.

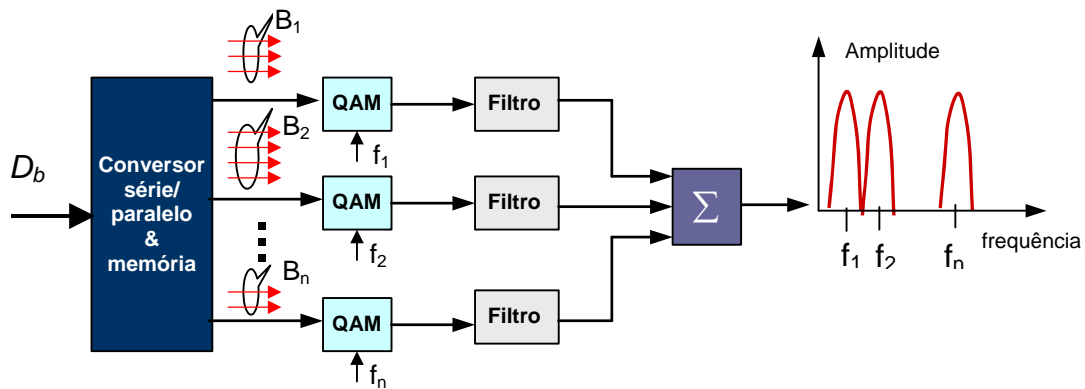



Figura 4.27 Esquema simplificado de um modulador DMT.

Por convenção os sub-blocos são ordenados de acordo com a dimensão da constelação que lhe corresponde. Os sub-blocos correspondentes às menores constelações são os primeiros a serem memorizados, enquanto os correspondentes às maiores constelações são menorizados em último lugar. No presença de constelações com a

mesma dimensão, são armazenados na memória em primeiro lugar os sub-blocos pertencentes às portadoras com o menor índice. Na figura 4.28 representa-se o processo de ordenamento descrito considerando uma trama com 22 bits, e um sistema com 6 sub-canais. De acordo com essa figura o tom 4 é modulado usando 4-QAM, os tons 0 e 5 usam 8-QAM, o tom 3 usa 16-QAM, e por fim os tons 1 e 2 usam 32-QAM. Note-se, ainda, que a duração de cada trama é igual a 250 μs, o que corresponde, portanto, a um débito de símbolo por sub-canal igual a 4 kbaud.

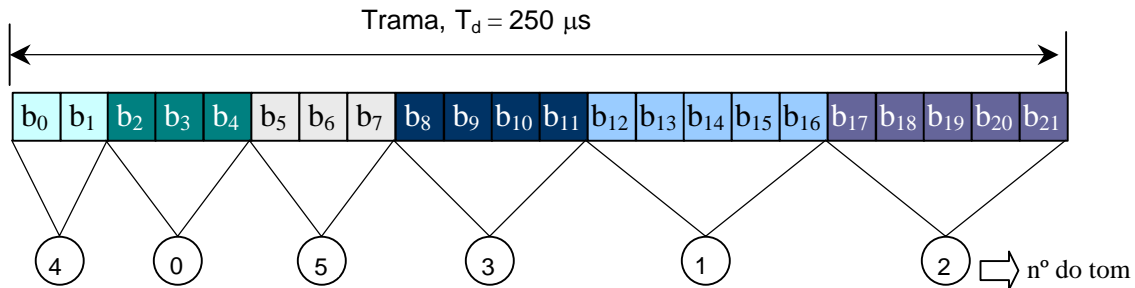



Figura 4.28 Atribuição do conteúdo da trama ADSL aos diferentes sub-canais.

Exemplo 4.4

A largura de banda de um canal passa banda é dada por $B_c=B_{min}(1+\alpha)$ onde α é o factor de excesso de banda, e B_{min} é igual ao débito de símbolo, ou seja $B_{min}=D_s$. Por sua vez, viu-se no Capítulo 3 que o débito binário se relaciona com o débito de símbolo através da relação $D_b=D_s\log_2 M$, onde M descreve o número de símbolos distintos da técnica de modulação. No caso específico do ADSL a largura de banda disponível é igual a 1104 kHz, e essa banda suporta 256 sub-canais, por isso a largura de banda de cada sub-canal é igual a $B_{s,c}$=4.3125 kHz. Como o débito de símbolo por sub-canal é de $D_{s,c}$=4 kbaud, a formulação anterior permite concluir que o factor de excesso de banda dos sub-canais é igual a α=0.078.

Exemplo 4.5

No caso específico do CE-ADSL a banda correspondente ao sentido descendente estende-se de 25.875 kHz até 1104 kHz e a banda correspondente ao sentido ascendente estende-se de 25.875 kHz até 138 kHz. Deste modo o número de sub-canais disponíveis no sentido descendente é igual a N_c= (1104-25.875)/4.3125=250. Deste sub-canais há 2 que são usados para funções de sincronismo, e por isso têm-se 248 canais úteis. O débito binário máximo suportado por este sistema é de 8 Mbit/s. Se todos os canais úteis estiverem activos, tal corresponde a um débito por sub-canal igual a $D_{b,c}$=32.258 bit/s. Tendo presente que $D_{s,c}$=4 kbaud, e atendendo à relação $D_{b,c}=D_{s,c}\log_2 M$, vem que M=267.7, pelo que a modulação a usar deverá ser 512-QAM (M=512). O número da canais no sentido ascendente é dado por N_c= (138-25.875)/4.3125=26. Neste caso também há dois canais que não são usados e por isso o número de canais úteis é igual a 24. Considerando um débito máximo no sentido ascendente igual a 800 kbit/s o débito por sub-canal é de $D_{b,c}$=33.333 bit/s, exigindo por isso também uma modulação 512-QAM.

Nos sistemas práticos a modulação DMT é implementada, de modo diferente do esquematizado na Figura 4.27, fazendo para isso uso da transformada de Fourier discreta inversa. Como o suporte de transmissão é muito adverso os blocos ATU-C e ATU-R usam também códigos correctores de erros de modo a reduzir a probabilidade de erro no receptor. Os códigos, normalmente, usados são os códigos convolucionais e os códigos de Reed-Solomon.

Os avanços na tecnologia ADSL conduziram a uma nova versão o ADSL2. Com esta nova versão o débito binário no sentido descendente pode atingir um máximo de 12 Mb/s e o débito ascendente de 1.2 Mb/s. Estes ganhos são conseguidos mantendo o plano de frequências típico do ADSL e melhorando as técnicas de modulação e codificação. Uma versão ulterior designada por ADSL2$^+$, permite, ainda, duplicar esses débitos, mas neste caso o plano de frequências é alterado e a largura de banda disponível aumenta para os 2.2 MHz.

4.5.5 VDSL

O VDSL (*Very-High Data Rate* DSL) é uma extensão do ADSL tendo surgido com o objectivo de permitir débitos de várias dezenas de megabits por segundo sobre os pares simétricos. Contudo, como se pode concluir da Tabela 4.3, estes débitos só são atingíveis para curtas distâncias, não permitindo em geral ir além de 1 km. Este facto faz com que esta tecnologia não possa ser considerada uma tecnologia de acesso autónoma, tendo de ser usada como um complemento de outras soluções. Essas soluções são, normalmente, baseadas na utilização de fibra óptica para substituir os cabos de alimentação e parte dos cabos de distribuição metálicos (ver Figura 4.2). A maior ou menor profundidade de penetração da fibra óptica na rede de acesso permite definir diferentes soluções (ver Secção 4.6), sendo a FTTC[53] e a FTTB[54], aquelas que habitualmente aparecem associadas ao VDSL. A primeira solução consiste em levar a fibra óptica até um ponto no passeio, onde alimenta um equipamento designado por ONU (*optical network unit*), enquanto a segunda solução aproxima um pouco mais a fibra óptica do subscritor, já que a ONU está situada no interior do edifício.

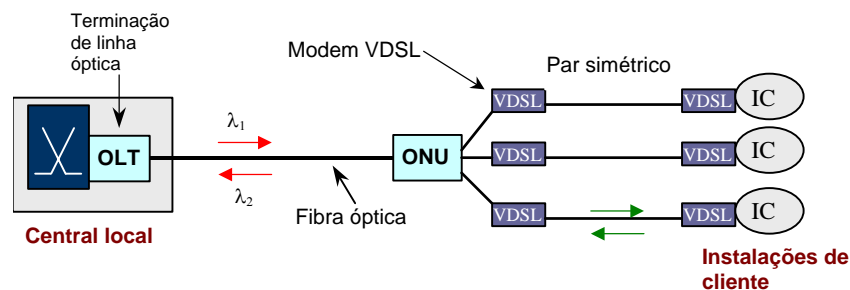



Figura 4.29 Utilização do VDSL como complemento de uma solução FTTx.

A Figura 4.29 exemplifica as soluções descritas. Na central local os sinais destinados a todos os utilizadores servidos por uma ONU são agregados usando uma técnica de multiplexagem apropriada e convertidos para o domínio óptico usando um equipamento

[53] *Fiber To The Curb*
[54] *Fiber To The Building*

designado por OLT (*optical line termination*). A bidireccionalidade no domínio óptico pode ser garantida usando uma técnica de duplexagem por divisão no comprimento de onda[55]. O sinal óptico é transmitido através de fibra óptica até à ONU, onde o processo é revertido. Os diferentes sinais depois de desmultiplexados vão alimentar um banco de modems VDSL, os quais estão interligados aos modems presentes nas instalações de cliente através de pares simétricos. O número de subscritores servidos por cada ONU é variável dependendo da maior ou menor penetração da fibra óptica, mas podem-se apontar valores típicos situados entre os 10 e 100 subscritores [GR05].

Enquanto os modems ADSL operam unicamente em modo assimétrico, os modems VDSL podem operar em modo assimétrico, ou simétrico, dependendo do tipo de aplicação requerida[56]. As tabelas 4.4 e 4.5 ilustram os débitos binários e os alcances das diferentes soluções.

Tabela 4.4 Alcance dos sistemas VDSL simétricos

Débito (Mb/s) Sentido descendente/ascendente	Alcance (km)
25	0.3
13	1
6.5	1.5

Tabela 4.5 Alcance dos sistemas VDSL assimétricos

Débito (Mb/s) Sentido descendente	Débito (Mb/s) Sentido ascendente	Alcance (km)
54	6.4	0.3
26	3.2	1
13	1.6	1.5

A duplexagem no âmbito dos sistemas VDSL pode ser obtida do mesmo modo que nos sistemas ADSL. A ocupação das bandas nos sistemas DDF está representada na Figura 4.30 para o caso do VDSL assimétrico. Como se pode ver a sub-banda situada entre os 300 e os 700 kHz é destinada ao tráfego ascendente. A sub-banda destinada ao tráfego descendente é delimitada inferiormente por 1 MHz e superiormente por um valor que pode variar entre f_{min}=10 MHz e f_{max}= 30 MHz. A sub-banda entre 0 e 120 kHz é destinada aos serviços telefónicos analógicos e aos serviços RDIS. A duplexagem baseada no CE como é de esperar requer a sobreposição das sub-bandas associadas ao tráfego nos dois sentidos. Esta sobreposição faz com que o efeito da paradiafonia (que é desprezável nos sistemas DDF) contribua para degradar fortemente o desempenho dos sistemas EC devido às levadas frequências usadas. Como no estado actual da tecnologia os canceladores de ecos só podem ser aplicados até frequências da ordem de 1 MHz, tal implica que a técnica de duplexagem que faz uso destes dispositivos não possa ser aplicada no VDSL simétrico.

[55] Usando comprimentos de onda diferentes para a transmissão nos dois sentidos.
[56] O modo assimétrico é usado para aplicações residenciais e o simétrico empresariais.

Quanto às técnicas de modulação a DMT é também a técnica mais usada no contexto dos sistemas VDSL. Uma outra técnica desenvolvida no contexto destes sistemas foi a SDMT (*Synchronized* DMT). A principal novidade associada a esta técnica consiste na utilização do TCM para garantir a duplexagem dos sinais.

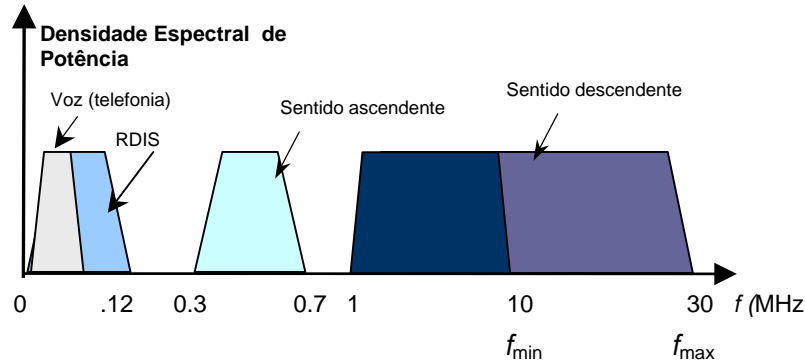



Figura 4.30 Ocupação da largura de banda de um sistema DDF-VDSL.

4.6 Soluções ópticas para a rede de acesso

4.6.1 Aspectos básicos

A introdução da fibra óptica na rede de acesso exige a presença de dispositivos com capacidade para realizar conversões electro-ópticas e óptico-eléctricas. Estes dispositivos situam-se, quer na central local (ou na unidade remota) enquadrados num bloco designado por OTN, quer num ponto situado mais ou menos próximo do subscritor e designado por ONU. A interligação entre a OLT e a ONU é feita em fibra óptica. A localização da ONU define diferentes soluções. A solução mais simples designa-se por FTTH (*Fiber To The Home*) e consiste em equipar as instalações de cada utilizador com uma ONU, o que implica levar a fibra óptica até casa do cliente. Esta solução apresenta, contudo, a desvantagem de ser muito cara e por isso antes de ser implementada convém explorar outra soluções alternativas, como por exemplo a FTTB (*Fiber To The Building*) e a FTTC (*Fiber To The Curb*). A primeira destas soluções consiste em instalar uma ONU por edifício e usar no interior do edifício uma infra-estrutura metálica para ligar os utilizadores à ONU, baseada por exemplo em pares simétricos, requerendo-se neste caso a utilização de modems VDSL nas duas extremidades do par. Por conseguinte, o sinal óptico correspondente ao fluxo descendente transmitido pela OLT é convertido para o domínio eléctrico na ONU e fluxo de dados gerado é desmultiplexado e entregue aos múltiplos modems VDSL Note-se que, neste caso os DSLAMs têm de ser deslocados para junto da ONUs e por isso os edifícios equipados com esta solução terão de dispor de uma “casa de telecomunicações” com capacidade para acomodar todo esta diversidade de equipamento, proporcionando ao mesmo tempo uma alimentação eléctrica fiável, o que normalmente requer a utilização de um banco de baterias eléctricas[57]. Uma alternativa mais económica, designada por FTTC, consiste em deslocar a ONU para um ponto de

---

[57] Nesta solução o plano de gestão do operador terá de penetrar profundamente no acesso de modo a poder gerir todo o equipamento presente na “casa das telecomunicações”.

distribuição na rede de acesso e usar a infra-estrutura de pares simétricos preexistente para interligar as instalações dos clientes às ONUs.

O modo como as diferentes ONUs estão interligadas à OLT define a arquitectura da componente óptica da rede de acesso. Essa interligação pode assumir uma das seguintes formas:

- Ligação ponto-a-ponto;
- Rede óptica passiva (PON[58]);
- Rede em anel.

Na ligação ponto-a-ponto tem-se uma fibra dedicada para interligar cada ONU, o que corresponde à implementação de uma topologia física em estrela. Na rede óptica passiva a fibra óptica é partilhada por vários utilizadores (16-32) usando para isso um repartidor de potência. A presença deste dispositivo vai fazer com que a PON seja um meio de transmissão partilhado entre múltiplas ONUs (topologia física em árvore), e por isso os sinais enviados por estas estão sujeitos a colisão, tal como acontece com as redes HFC estudadas no Capítulo 1. Para evitar essas colisões as PONs requerem a utilização de algoritmos apropriados para regular o processo de comunicação, o que faz com que os protocolos usados nas redes PON sejam muito mais complexos, do que aqueles que são usados nas redes baseadas na ligação ponto-a-ponto. Será de destacar que em ambas as arquitecturas a bidireccionalidade da comunicação é garantida usando duplexagem por divisão do comprimento de onda cuja definição foi apresentada na secção anterior.

A rede em anel, como o próprio nome indica, interliga as diferentes ONUs à OLT usando uma topologia física em anel. Contrariamente às soluções anteriores, neste caso um caminho entre a OLT e uma ONU pode requerer diferentes conversões óptico-eléctricas. Normalmente, esta arquitectura é implementada recorrendo à tecnologia SDH que irá ser estudada no capítulo 5.

4.6.2 PON

A estrutura de uma PON está representada na Figura 4.31. O sinal óptico transmitido pela OLT é repartido pelas diferentes ONUs usando um repartidor óptico passivo. Embora a bidireccionalidade possa ser garantida usando duas fibras, uma para cada sentido, a maioria das implementações usa uma única fibra e duplexagem por divisão no comprimento de onda como se mostra na Figura 4.31. Esta técnica de demultiplexagem requer a utilização de um dispositivo óptico adicional em cada ONU para fazer a multiplexagem/desmutiplexagem dos dois comprimentos de onda usados[59]. A figura indica os comprimentos de onda usados nas PONs de primeira geração designadas por APON. Nestas redes a conectividade era garantida usando ATM, ou seja o fluxo de informação é segmentado em células, as quais são entregues à ONU apropriada de acordo com o seu endereço de destino contido no cabeçalho da célula. A comunicação ascendente requer a utilização de um protocolo MAC[60] apropriado para ultrapassar a limitação associada ao facto do meio ser partilhado. Para além disso, o receptor óptico

[58] *Passive Optical Network*
[59] Embora não representado na figura a OLT terá de usar um dispositivo idêntico.
[60] *Medium Access Control*

usado na OLT tem de ser concebido para receber rajadas de bits (em vez de sequências continuas como acontece nos receptores convencionais) e além disso para receber sinais com diferentes amplitudes devido ao facto do espaçamento dos diferentes ONUs à OLT ser variável. Como consequência o limiar de decisão terá de ser ajustável de rajada, para rajada, requerendo por isso a presença de um preâmbulo associado a cada rajada.

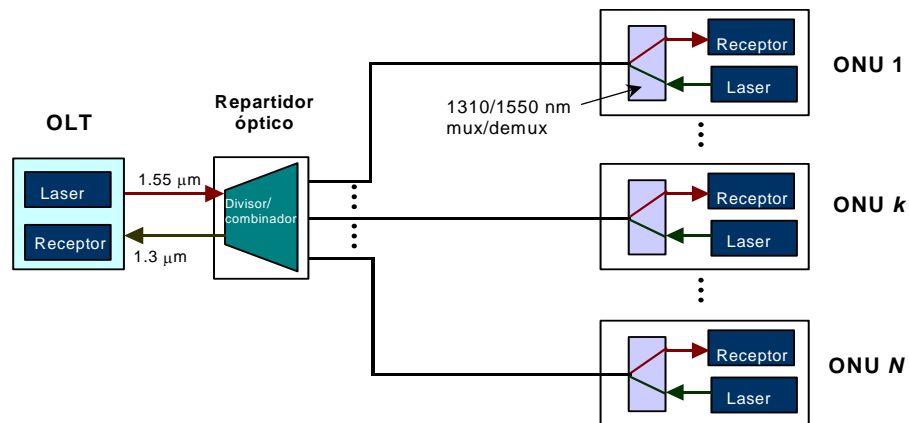



Figura 4.31 Estrutura de uma rede óptica passiva.

As novas gerações de PONs fundamentalmente dividem-se em duas grandes famílias: a GPON (*Gibabit PON*) e a EPON (*Ethernet* PON). A primeira resulta da evolução das APONs e surporta, quer células, quer tramas. A segunda, como o próprio nome indica, suporta unicamente tramas Ethernet. A norma GPON define um factor de repartição máximo de 64 e um débito binário máximo de transmissão de cerca de 2.5 Gbit/s. Por sua vez, a norma EPON define um débito de transmissão de 1.25 Gbit/s e usa um factor de repartição típico de 16. Neste caso, na pior das situações (todas as ONUs activas), uma EPON consegue garantir um débito de cerca de 80 Mbit/s a cada ONU. Na realidade os receptores dos sistemas EPON são projectados para um débito de 100 Mbit/s, o qual é garantido na maior parte do tempo, já que a probabilidade de todas as ONUs estarem simultaneamente activas é muito reduzida. No caso em que usa FTTH, como já se referiu a ONU está situada nas instalações de cliente, e por isso o débito disponibilizado a cada utilizador, quando se usa uma rede EPON é de 100 Mbit/s.

4.7 Referências

[Bi00] J. A. C. Bingham, *ADSL, VDSL, and MultiCarrier Modulation*, John Wiley&Sons, 2000.

[Co98] G. J. Cook, “Transmission planning”, *Telecommunications Engineer`s Reference Book,* Focal Press, pág. 40/1-40/13, 1998.

[Ga03] M. Gagnaire, *Broadband Local Loops for High-Speed Internet Access*, Artech House, Inc., 2003.

[GR05] P. E. Green Jr., *Fiber To The Home-The New Empowerment*, John Wiley&Sons, 2006.

[Gr04] F. M. Groom, K. M. Groom, *The Basics of Voice over Internet Protocol,* International Engineering Consortium, 2004.

[Ma96] K. Maxwell, "Asymmetric Digital Subscriber Line: Interim Technology for the Next Forty Years," *IEEE Com. Magazine*, pág. 100-106, Outubro 1996.

[Nu92] M. S. Nunes e A.J. Casaca, *Redes Digitais com Integração de Serviços*, Editorial Presença, 1992.

[Re95] W. D. Reeve, *Subscriber Loop Signaling and Transmission Handbook*, IEEE Press 1995.

5 Tecnologias de transporte

5.1 Princípios do FDM e do TDM

Como geralmente o custo da via de transmissão (inclui meio de transmissão e repetidores) é elevado, é conveniente aproveitar a largura de banda disponível para transmitir pela mesma via mais do que um canal telefónico. Os sistemas usados para este efeito são designados por sistemas de telefonia múltipla (ou multiplexer).

Existem essencialmente duas técnicas distintas de realização de telefonia múltipla: a divisão na frequência (FDM, frequency division multiplexing) e a divisão no tempo (TDM, time division multiplexing). Nos sistemas de transmissão em fibra óptica é usual falar-se em multiplexagem por divisão do comprimento de onda (WDM, wavelength division multiplexing), para designar os sistemas FDM.

5.1.1 Multiplexagem por Divisão na Frequência

A multiplexagem FDM baseia-se no princípio de que o espectro de um sinal modulado em amplitude é exactamente o espectro do sinal em banda base transladado para a banda de frequências centrada na frequência da portadora. Assim, para transportar na mesma via de transmissão diferentes sinais, basta que eles se encontram modulados por portadoras com frequências tais que as diferentes réplicas espectrais não interfiram entre si.

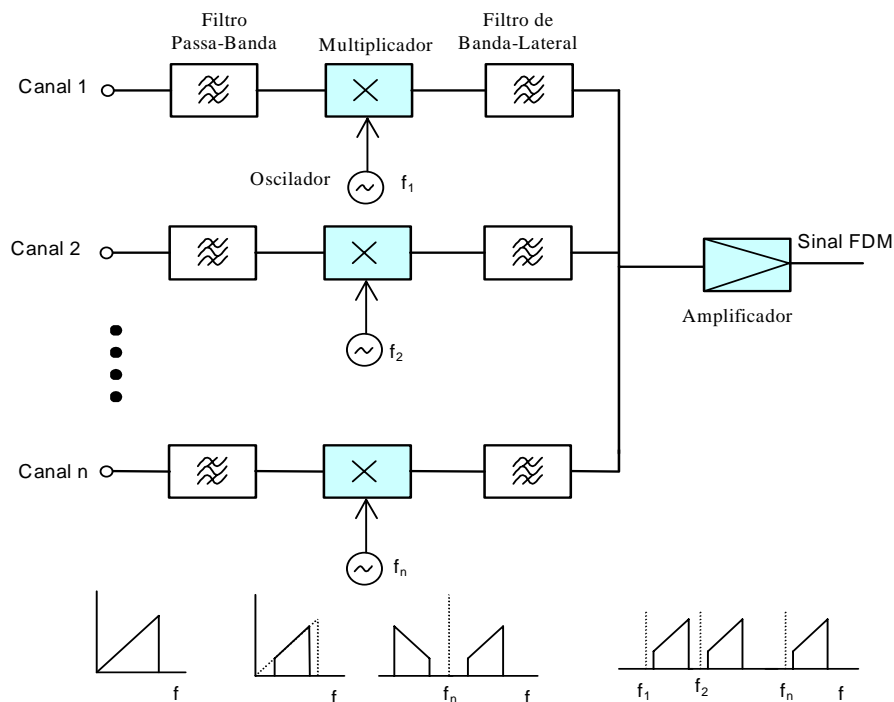



Figura 5.1 Esquema de blocos de um multiplexador FDM.

Na Figura 5.1 apresenta-se o esquema de blocos de um multiplexador FDM. Os sinais de áudio presentes na entrada são filtrados de modo a garantir que o espectro desses sinais não excede a banda entre os 300 e os 3400 Hz. Cada um dos sinais de áudio filtrado é em seguida multiplicado (usando um modulador equilibrado) por uma portadora com uma determinada frequência. Tem-se, assim, uma modulação de banda lateral dupla com supressão de portadora.

Para evitar interferência entre os diferentes canais essas portadoras estão distanciadas de 4 kHz. A fase seguinte do processo consiste em eliminar uma das bandas (no esquema da Fig. 5.1 eliminou-se a inferior) usando um filtro passa-banda, designado por filtro de banda lateral. Os diferentes sinais resultantes desse tratamento são em seguida combinados de modo a originar um sinal FDM.

A operação de desmultiplexagem é realizada multiplicando cada um dos canais por uma portadora com a mesma frequência e fase que a usada na multiplexagem, seguida de filtragem passa-baixo.

A multiplexagem de um número elevado de canais telefónicos não é efectuada através de uma única operação de modulação e filtragem, mas antes, através de etapas sucessivas, formando uma estrutura hierárquica. O número de canais presentes em cada nível hierárquico é fruto de normalização do ITU-T. Na Tabela 5.1 indica-se a designação de cada nível, o número de canais e a banda ocupada.

Tabela 5.1 Hierarquia dos sistemas FDM

Nome	*Número de canais*	*Banda ocupada*
Grupo	12	60-108 kHz
Supergrupo	60	312-552 kHz
Grupo mestre	300	812-2044 kHz
Super grupo mestre	900	8.516-12.338 MHz

Normalmente, para um melhor aproveitamento das capacidades dos meios de transmissão usados no contexto da transmissão FDM (cabos coaxiais e feixes hertzianos), os diferentes grupos eram associados de modo apropriado. Por exemplo, a recomendação G 333 do ITU-T, destinada à transmissão sobre cabo coaxial, apontava para a associação de 12 super grupos mestre, o que corresponde a 10 800 canais telefónicos e uma banda ocupada entre os 4.332 e 59.684 MHz.

Para terminar estas breves considerações sobre o FDM será de destacar que as redes telefónicas modernas já praticamente não fazem uso desta técnica de multiplexagem. No entanto, o FDM continua a ser a técnica dominante nas redes de distribuição de televisão por cabo, já que a transmissão analógica continua a ser predominante nesta área. No caso das redes híbridas (fibra/coaxial) o FDM aparece vulgarmente com a designação de multiplexagem de sub-portadora (*subcarrier multiplexing*), o que se explica pelo facto de a portadora principal nessas redes ser uma portadora óptica, funcionando as portadoras eléctricas usadas para obter o sinal FDM como sub-portadoras.

5.1.2 Multiplexagem por divisão no comprimento de onda

A multiplexagem por divisão no comprimento de onda permite explorar de modo eficiente a largura de banda das fibras ópticas, juntando numa mesma fibra um certo número de portadoras ópticas, cada uma com o seu comprimento de onda. O conceito de

multiplexagem por divisão no comprimento de onda ou WDM (*wavelength division multiplexing*) está ilustrado na Figura 5.2.

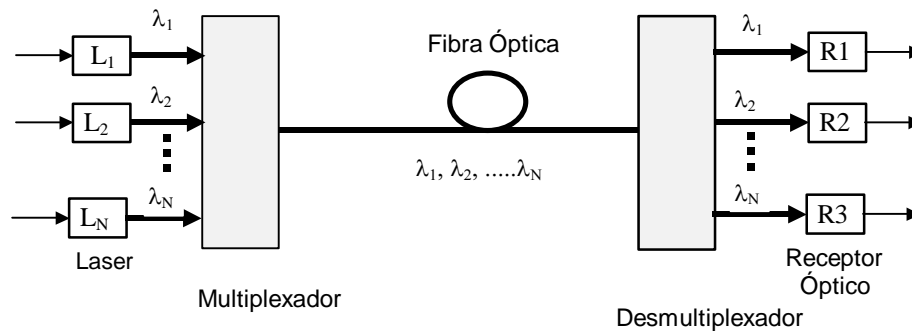



Figura 5.2 Processo de multiplexagem por divisão no comprimento de onda.

Antes da operação de multiplexagem é necessário converter os N sinais digitais a serem agregados para o domínio óptico. Para isso, esses sinais modulam (directamente ou externamente) N lasers de semicondutor, onde cada laser emite no seu próprio comprimento de onda. Os sinais ópticos resultantes são em seguida multiplexados no domínio óptico, gerando um sinal WDM, que é injectado numa fibra óptica monomodal. Na outra extremidade da fibra os diferentes comprimentos de onda são separados (desmultiplexados) e os correspondentes sinais detectados através de receptores ópticos. Estes receptores são responsáveis por regenerar os sinais e recuperá-los para o domínio eléctrico.

A tecnologia actual permite modular os lasers com débitos binários que podem ir até aos 10 Gbit/s, e permite multiplexar até 160 portadoras ópticas numa única fibra óptica. O espaçamento entre os diferentes comprimentos de onda depende da tecnologia de multiplexagem usada, podendo-se atingir espaçamentos da ordem dos 25 GHz. Não é propósito deste curso estudar as tecnologias usadas para implementar o WDM, mas não se pode deixar de referir que o sistema de multiplexagem/desmultiplexagem de comprimento de onda mais simples é constituído pelo prisma óptico, usado na decomposição da luz branca.

5.1.3 Multiplexagem por divisão no tempo

De acordo com o teorema da amostragem um sinal banda-base com largura de banda B (Hz), pode ser univocamente determinado a partir das suas amostras, desde que estas sejam tomadas em intervalos de tempo uniformemente espaçados de $T_a \leq 1/(2B)$. Como o sinal amostrado está em estado desligado uma parte significativa do tempo, pode-se aproveitar esses intervalos sem sinal, para transmitir as amostras correspondentes a outros sinais. De forma resumida é este o princípio do TDM. Este princípio está ilustrado na Figura 5.3. Segundo essa figura os sinais passa-baixo correspondentes aos diferentes canais são amostrados usando portas lógicas, que são activadas num curto intervalo de tempo, pela acção das sequências de pulsos representadas na Figura 5.3(b). Todas as sequências têm a mesma frequência de repetição, mas estão desfasadas no tempo. O sinal

TDM resulta da interposição das sequências de pulsos moduladas pelos sinais provenientes dos diferentes canais.

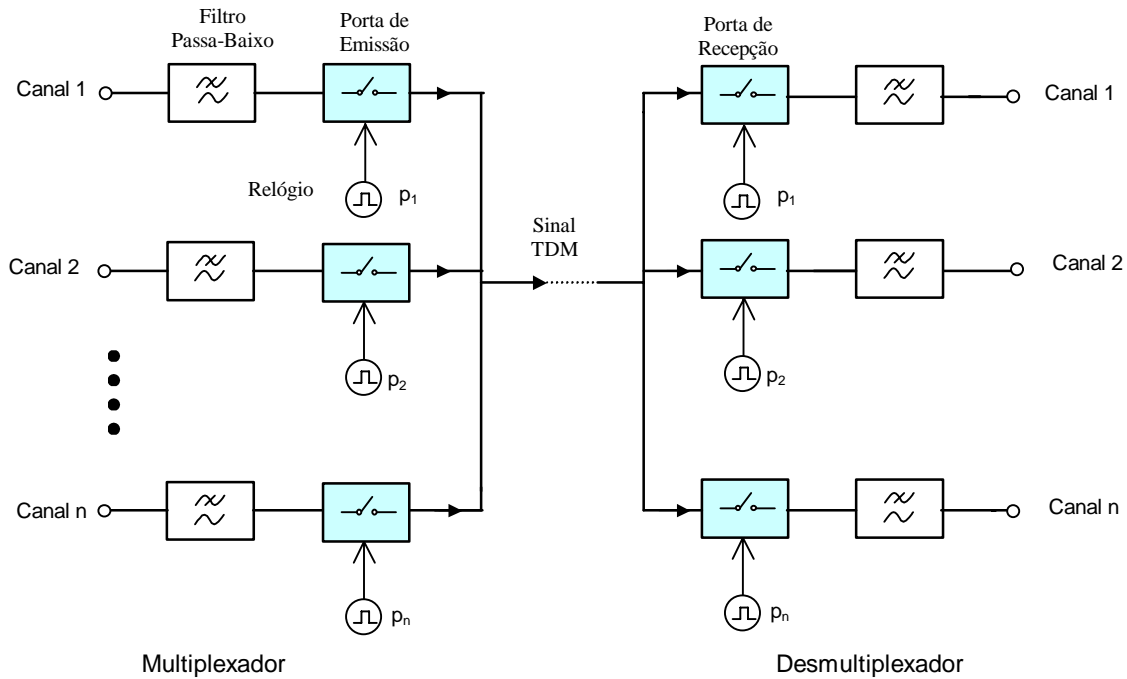
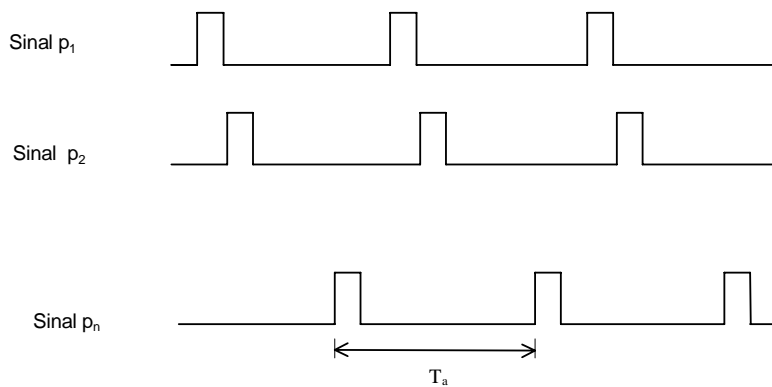



Figura 5.3 Princípio do TDM.

No desmultiplexador as portas são activadas por sequências de pulsos sincronizadas com as usadas no multiplexador. Assim, para além dos pulsos correspondentes aos canais de informação é necessário transmitir um sinal apropriado para sincronizar os pulsos responsáveis pelo controlo das portas lógicas do multiplexador e do desmultiplexador. O sinal transmitido durante um período de repetição T_a é constituído por um determinado número de hiatos temporais (*time-slots*) sendo, por exemplo, um destinado ao sinal de

sincronização e os outros às amostras dos diferentes canais. O sinal completo é designado por **trama** e o sinal de sincronismo por **sinal de enquadramento de trama**.

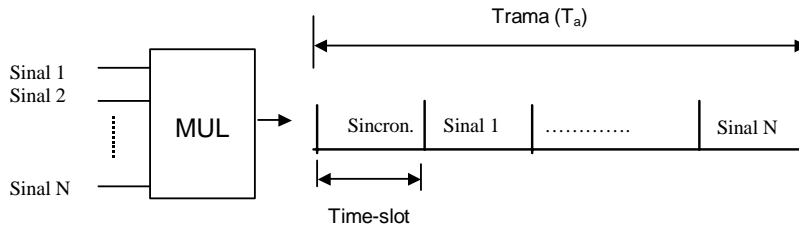



Figura 5.4 Estrutura da trama TDM.

O sistema TDM representado na Figura 5.3 usa modulação de pulsos em amplitude (PAM, *pulse amplitude modulation*) daí designar-se por TDM-PAM. Este método de modulação é, como se sabe, analógico, não sendo por isso apropriado para transmissão a grandes distâncias. Como consequência, é usado na generalidade dos casos PCM para a transmissão do TDM (TDM-PCM).

Têm-se duas técnicas para obter um sinal TDM-PCM. A primeira consiste em multiplexar as amostras analógicas seguida de codificação. A segunda realiza primeiro a codificação e em seguida a multiplexagem dos bits correspondentes. No primeiro caso, a multiplexagem é analógica e usa-se um *codec* comum para todos os canais telefónicos, para realizar a digitalização do sinal multiplexer. No segundo caso, a multiplexagem é digital, havendo necessidade de usar um *codec* por cada canal telefónico. Normalmente, nos sistemas telefónicos usa-se o segundo método, ou seja, multiplexagem digital.

Existem dois métodos para entrelaçar a informação relativa aos diversos canais telefónicos: **interposição de bit** e **interposição de palavra**. No primeiro caso, a cada canal atribui-se um *time slot* constituído por um único bit, enquanto no segundo caso atribui-se um *time-slot* mais longo, constituído por vários bits (palavra). No caso particular em que a palavra é constituída por 8 bits o entrelaçamento designa-se por interposição de octeto.

5.1.3.1 Sincronização de elementos de rede

Os multiplexadores TDM, assim como, a generalidade dos elementos das redes de telecomunicações digitais, requerem uma fonte de temporização, ou por outras palavras um relógio, de modo a sincronizar todas as operações realizadas por esses elementos de rede. Um sinal de temporização é um sinal gerado por um relógio e é caracterizado por ser um sinal periódico de período T e frequência $f_0 = 1/T$. As formas típicas usadas para descrever são ondas sinusoidais e quadradas, como se representa na Figura 5.5. Um sinal de relógio, ainda pode ser modelado por uma sequência de pulsos espaçados de T e localizados em instantes significativos, ou seja nos instantes em que o sinal de temporização influência o processo que controla. Na figura 5.5, a sequência de instantes significativos é constituída pelos instantes de passagem por zero no flanco ascendente dos sinais representados. Os sinais em que o intervalo de tempo entre instantes significativos tem pelo menos em média a mesma duração, ou as mesmas durações no caso desses intervalos serem múltiplos de um mais pequeno, designam-se por sinais **isócronos.**

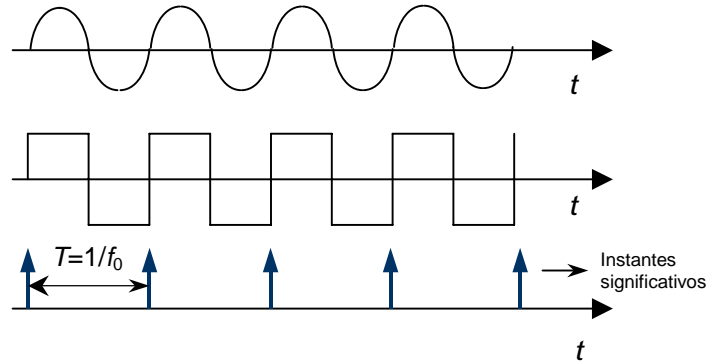



Figura 5.5 Sinais de temporização e identificação dos instantes significativos.

Os relógios reais são concebidos para operarem a uma determinada frequência nominal (f_0). Na prática, contudo, devido a limitações físicas o relógio opera a uma frequência real (f_r), que se pode afastar mais ou menos da frequência nominal. O parâmetro que contabiliza o desvio da frequência real de uma relógio relativamente à sua frequência nominal designa-se por precisão de relógio e é definido por

$$\text{Precisão} = \frac{|f_r - f_0|}{f_0} \tag{5.1}$$

sendo expresso usualmente em p.p.m (partes por milhão). Tendo em conta a precisão, define-se usualmente uma hierarquia com quatro níveis (*stratum)*, com os relógios com precisão mais elevada (relógios atómicos) pertencendo ao *stratum* 1.

Tabela 5.2 Níveis de precisão dos relógios usados no sincronismo de redes

Nível	*Stratum*1	*Stratum*2	*Stratum*3	*Stratum*4
Precisão	1×10^{-11}	1.6×10^{-8}	4.6×10^{-6}	3.2×10^{-5}

Outro aspecto a ter em consideração quando se analisa a problemática da sincronização de elementos de rede é relação de temporização entre sinais. Dois sinais isócronos dizem-se **síncronos** se tiverem a exactamente a mesma frequência e o seu desvio de fase for constante. Por sua vez, dois sinais isócronos dizem-se **plésiocronos** se as suas frequências nominais forem idênticas, mas as frequências reais são diferentes. Os sinais gerados por dois relógios oscilando independentemente são sempre plésiocronos. Por isso, para gerar sinais síncronos é necessário que todos os relógios sejam sincronizados a partir de um relógio central. No caso específico das redes de telecomunicações este relógio central é normalmente um relógio atómico (césio ou rubídio)[61] e designa-se por referência primária ou PRC (*Primary Reference Clock)*. As redes em que esta estratégia é seguida, designam-se por redes síncronas, e requerem a existência de uma rede de sincronização,

[61] Também se pode usar o sistema GPS (*Global Positioning Satellite System*)

apropriada para distribuir o sinal de relógio[62] proveniente da referência primária a todos os elementos de rede (ver Figura 5.6).

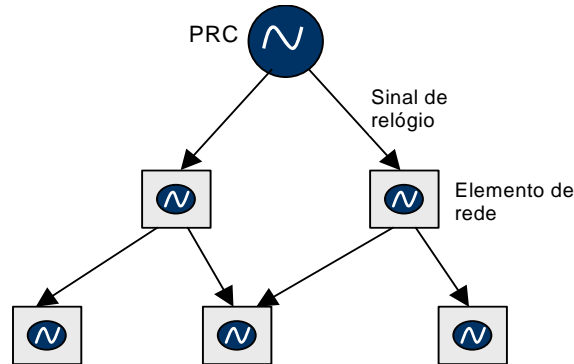



Figura 5.6 Exemplo de uma rede síncrona.

5.1.3.2 Sistemas de multiplexagem primários

Nos sistemas de multiplexagem primários (TDM-PCM) a sequência binária transmitida é estruturada em tramas de duração igual a 125 μs (correspondendo a uma frequência de amostragem de 8 kamostras/s). Cada trama contem uma amostra codificada de cada um dos canais de voz multiplexados, juntamente com bits adicionais para funções de sinalização e de sincronismo de trama.

A nível internacional são usados, normalmente, dois sistemas de multiplexagem primários, que têm a particularidade de serem incompatíveis entre si: o sistema europeu designado por CEPT1, definido pela *Conference of European Post and Telecommunications* (CEPT), e o sistema americano designado por DS1 (*Digital Signal*-1). Ambos foram fruto de normalização por parte do ITU-T (ver tabela 5.3).

O sistema CEPT1, hoje designado usalmente por E1, suporta 30 canais telefónicos, enquanto o DS1 foi projectado para 24. Ambos usam uma frequência de amostragem de 8 kamostras/s e 8 bits por amostra, mas enquanto o CEPT1 usa a lei de compressão A, o DS1 usa a lei de compressão μ.

Tabela 5.3 Sistemas de multiplexagem primários europeu e americano.

Parâmetros	Sistema europeu (Recomendação G 732)	Sistema americano (Recomendação G 733)
Frequência de amostragem	8 kHz	8 kHz
Número de níveis de quantificação	256	256
Número de bits por amostra	8	8
Débito binário por canal	64 kbit/s	64 kbit/s
Lei de quantificação	A (=87.6)	μ (=255)
Característica de compressão	13 segmentos	15 segmentos

---

[62] Este sinal de relógio é usualmente um sinal a 2.048 MHz (analógico), ou a 2.048 Mbit/s (digital).

Número de *time-slots*	32	24
Número de canais telefónicos	30	24
Número de bits por trama	32x8=256	24x8+1=193
Débito binário total	256x8 kHz=2.048 Mbit/s	193x8 kHz=1.544 Mbit/s
Enquadramento de trama	em bloco *palavra de 7 bits no* time slot *0 das tramas ímpares*	distribuído sequência 101010… constituída pelo 193 ésimo bit das tramas ímpares
Sinalização	*time slot* 16 à razão de 4 bits por canal repartidos por 16 tramas (multitrama)	oitavo bit de cada canal numa trama em cada seis

A estrutura da trama de um sistema primário de 30 canais está representada na Figura 5.7. A trama é dividida em 32 *time-slots,* cada um com 8 dígitos binários (octetos), de modo que o débito binário total é de 8x8x32 kbits/s=2.048 Mbit/s. O *time slot* 0, das tramas ímpares, é destinado ao **padrão de enquadramento de trama** (PET) e o das tramas pares à transmissão de um alarme. Os *time slots* de 1-15 e de 17-31 são destinados aos canais telefónicos de voz. O *time-slot* 16 é alocado para a sinalização dos canais de voz. A sinalização é transmitida canal a canal à razão de 4 bits por canal, possibilitando, assim, a transmissão da informação de sinalização de dois canais em cada trama. São necessárias, deste modo, 15 tramas para transmitir a informação de sinalização dos 30 canais. A essas tramas é adicionada uma outra de modo a formar uma **multitrama** constituída por 16 tramas. O padrão de enquadramento de multitrama situa-se no *time-slot* 16 da décima-sexta trama. Em síntese, neste sistema todos os canais de voz têm à sua disposição para sinalização 4 bits de 2 em 2 ms (16x125μs). Como se irá ver posteriormente, esta técnica de sinalização corresponde à chamada sinalização de canal associado. No caso da sinalização em modo-comum (ex: sistema de sinalização nº7), o *time-slot* 16 é usado para proporcionar uma via de sinalização a 64 kbit/s, usada conforme as necessidades por todos os 30 canais.

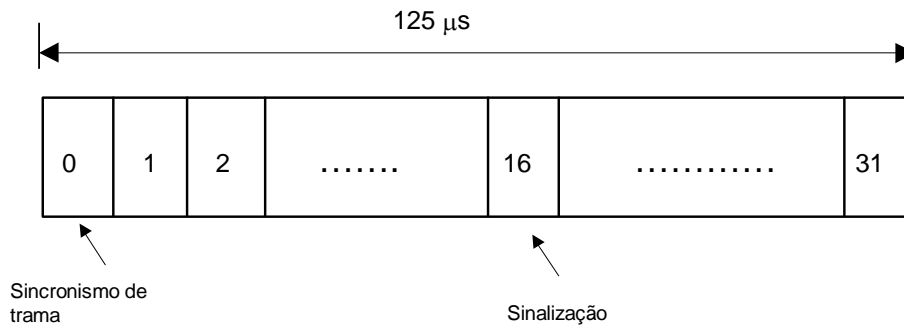



Figura 5.7 Estrutura da trama de um sistema TDM-PCM primário (E1) de 2.048 Mbit/s.

A trama básica de um sinal DS1 é constituída por 193 bits, a que correspondem os 24 *time-slots* dos canais de voz (cada um com 8 bits), mais um bit adicional (bit *F*) para funções de enquadramento de trama. O débito binário total é assim igual a 193×8 kbit/s=1.544 Mbit/s.

Nas tramas ímpares o bit *F* representa o padrão 1,0,1,0,1,…, que é usado para fins de enquadramento de trama. Repare-se que este enquadramento é **distribuído** em alternativa ao **enquadramento em bloco** do sistema E1. A informação transportada no bit *F* das tramas pares corresponde ao padrão 001110,…., é usada para definir uma multitrama de 12 tramas. Na sexta e duodécima tramas de cada multitrama, o oitavo bit de cada *time-slot*, é usado para transmitir a informação de sinalização do correspondente canal. Nesse *time-slot* o canal de voz é codificado só com 7 bits, o que conduzirá a uma ligeira degradação do desempenho devido ao aumento do ruído de quantificação. Destaque-se que, neste sistema, para proporcionar uma via de sinalização a 64 kbit/s para operação em modo comum, é necessário sacrificar um dos canais de voz, ficando disponíveis só 23 canais para tráfego de voz.

5.1.3.3 Hierarquias de ordem superior

Os sistemas de multiplexagem primários de 30 e 24 canais são usados como bloco básico para a construção de sistemas multiplex de ordem superior, estruturados segundo um esquema hierárquico bem definido. Em cada nível hierárquico, diferentes sequências binárias conhecidas como **tributários** são combinadas usando um multiplexador. A saída desse multiplexador origina um sinal que funciona como tributário para o multiplexador de ordem seguinte da hierarquia.

Diz-se que a multiplexagem é **síncrona**, quando os diferentes tributários que servem de entrada para um multiplexador de ordem superior têm todos exactamente o mesmo débito binário e além disso estão em fase. Contudo, na primeira geração de sistemas de multiplexagem digitais de ordem superior os diferentes tributários não estão geralmente perfeitamente sincronizados. Embora, todos tenham o mesmo débito binário nominal, como as suas fontes de origem são controladas por relógios distintos e independentes, será de esperar ligeiras flutuações nesses débitos dentro dos limites impostos pela tolerância desses relógios. Esses tributários designam-se, por isso, **plesiócronos** (do grego plésio significa quase) e a hierarquia que lhes corresponde **hierarquia digital plesiócrona**.

Exemplo 5.1

Considere-se que o relógio usado para controlar um multiplexador primário de 30 canais tem uma precisão de 50 partes por milhão (ppm). Pretende-se determinar os limites da flutuação da frequência desse relógio.

Solução: Como se sabe, a frequência nominal desse relógio é igual a 2.048 MHz. Assim, a sua flutuação é dada por

$$\Delta f = \frac{50}{10^6} x 2.048\ \mathrm{MHz} = 102.4\ \mathrm{Hz}$$

Assim, a frequência desse relógio flutuará entre 2.047898 MHz e 2.048102 MHz.

A segunda geração dos sistemas de multiplexagem de ordem superior é já baseada em tributários síncronos e, por isso, designa-se por **hierarquia digital síncrona.**

5.2 Hierarquia digital plesiócrona

Existem três conjuntos de normas, incompatíveis entre si, para a formação da hierarquia digital plesiócrona (PDH, *plesiochronous digital hierarchy*): a europeia, a americana e a japonesa. As normas europeias são baseadas no sistema multiplico primário de 30 canais, enquanto as americanas e japonesas são baseadas no sistema de 24 canais. A hierarquia europeia está representada na Figura 5.8, enquanto a americana está descrita na tabela 5.4.

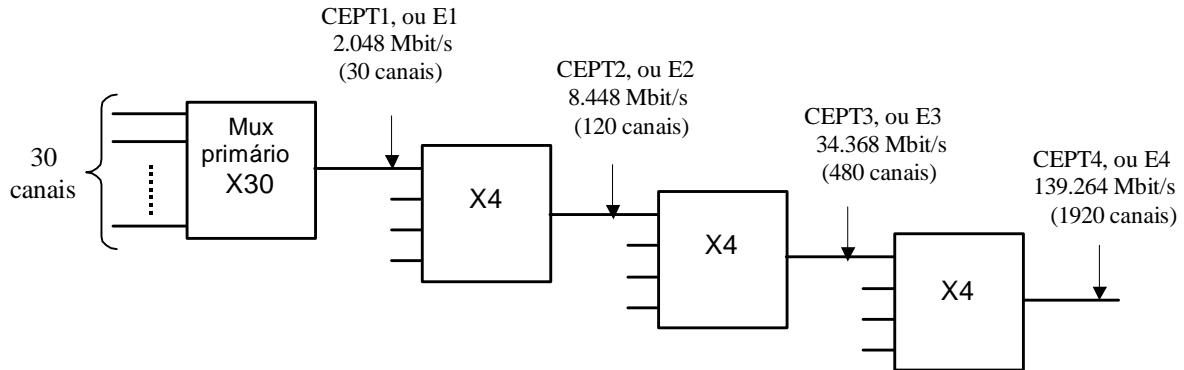



Figura 5.8 Hierarquia plesiócrona europeia.

Tabela 5.4 Hierarquia plesiócrona americana

Nível	Número de canais	Débito binário (Mbit/só)
DS-1	24	1.544
DS-2	96	6.312
DS-3	672	44.736
DÊ	4032	274.176

Todas as hierarquias plesiócronas de ordem superior usam interposição de bit. Nota-se, de imediato, que o débito binário do sinal resultante da multiplexagem de N tributários é superior ao débito binário obtido multiplicando por N os débitos binários de cada um dos tributários. Isto acontece porque é necessário inserir bits adicionais por duas razões. A primeira é o enquadramento de trama. Um desmultiplexador de ordem superior necessita (como no caso do primário) de reconhecer o início da trama, no sentido de encaminhar correctamente os bits recebidos para o tributário destinatário. Para isso, é necessário inserir uma sequência de bits predeterminada, designada por **padrão de enquadramento de trama** (PET). No caso da hierarquia europeia este padrão é inserido em bloco no início da trama, enquanto no caso da hierarquia americana é distribuído ao longo da trama.

A segunda razão que leva à introdução de bits adicionais prende-se com a necessidade de realizar a **justificação** de trama. Este processo tem como finalidade garantir uma operação correcta dos multiplexadores e desmultiplexadores, mesmo no caso em que o débito binário dos diferentes tributários flutua relativamente ao seu valor nominal.

5.2.1 Acomodação das flutuações dos tributários

A constituição de um sinal multiplex de ordem N requer que as tramas dos sinais dos tributários de ordem N-1 estejam perfeitamente sincronizadas, tanto em frequência como em fase. Como nos sistemas plesiócronos tal não acontece, será necessário antes da operação de multiplexagem propriamente dita submeter as sequências binárias provenientes dos diferentes tributários a um processamento apropriado. Para isso, usam-se memórias com capacidade para armazenar uma trama (de ordem N-1) completa (L bits), designada normalmente por **memória elástica.** Essa memória é escrita ao ritmo f_k, determinado pelo relógio extraído do canal k e é lida ao ritmo f_k', derivado a partir do relógio interno do multiplexador de ordem N. Destaque-se que f_k é igual ao débito binário D_k do canal k, enquanto f_k' é igual ao débito binário D_k', com que a informação é lida (ver Figura 5.9).

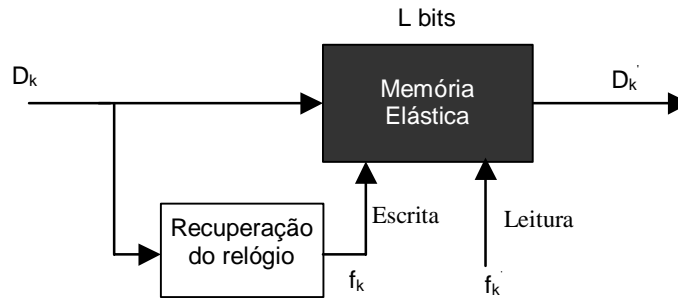



Figura 5.9 Estrutura de uma memória elástica.

Nos sistemas plesiócronos é pouco provável que f_k seja igual a f_k'. Então, se f_k é maior do que f_k' (o ritmo da sequência de entrada é superior ao ritmo da sequência de saída), o número de bits armazenados na memória (não lidos) tende a aumentar, podendo acontecer que uma nova trama seja escrita na memória antes da anterior ter sido lida. Nesse caso, esta última trama está, assim, definitivamente perdida. Na situação oposta (f_k menor que f_k'), como o número de bits armazenados (não lidos) tende a decrescer, a leitura poderá acontecer antes da escrita, sendo, por conseguinte, lida a trama precedente. Esta perda ou repetição de exactamente uma trama designa-se por *slip* controlado, ou simplesmente *slip,* e produz-se com um período dado por

$$T_s = \left|\frac{L}{D_k - D_k'}\right| = \left|\frac{L}{\Delta D_k}\right| \tag{5.2}$$

Exemplo 5.2
Pretende-se determinar a estabilidade relativa de dois relógios independentes para garantir um período de slips de 20 h, considerando o sistema de multiplexagem primário europeu.

Solução: Se se tiver em conta que a duração da trama é dada por $T_a = L/D_k$, a equação (5.2) permite escrever

$$\frac{\Delta D_k}{D_k} = \frac{T_a}{T_s}$$

Então, como T_a =125μs e T_s =20x60x60=72000 s, conclui-se que

$$\frac{\Delta D_k}{D_k} = 1.74x10^{-9}$$

ou seja, a estabilidade relativa dos relógios deverá ser de 1.74 partes em 10^9.

Em telefonia, os *slips* correspondem à perda ou repetição de uma amostra do sinal de voz e produzem impulsos parasitas audíveis (cliques), que são perfeitamente toleráveis desde que a sua frequência não seja muito elevada. Em transmissão de dados, ou voz encriptada, a perda de uma palavra de 8 bits pode ter consequências muito mais nefastas, sendo por isso conveniente tentar evitá-los.

Para evitar os *slips,* compensam-se as diferenças de débitos binários através de uma técnica designada por **justificação positiva.** Com esta técnica o ritmo do relógio de leitura f_k' é feito superior ao ritmo máximo do canal $f_k + \Delta f_k$, e para evitar o esvaziamento da memória é inserido de tempos em tempos na sequência binária de saída, um bit não portador de informação (**bit de justificação**). O processo de justificação é realizado controlando a ocupação da memória elástica. Sempre que esta desce abaixo de um certo limiar, é gerado um pedido para uma acção de justificação. Na próxima oportunidade (trama seguinte), o relógio de leitura é inibido durante um período de relógio, permitindo que um bit de justificação seja inserido num *time-slot* bem determinado. Antes da operação de inserção do bit de justificação, a localização do referido bit é codificada em bits existentes na trama para esse fim, que se designam por bits de **indicação de justificação**. Com base na informação transmitida pelos bits de indicação de justificação, os bits de justificação são removidos no desmultiplexador.

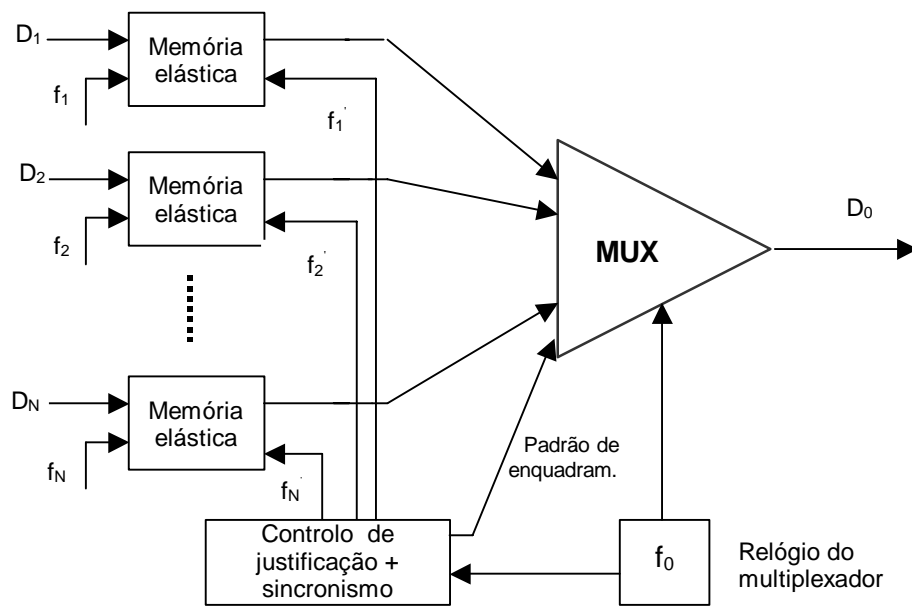



Figura 5.10 Estrutura de um multiplexador TDM plesiócrono.

Assim, a informação gerada a um débito superior ao nominal pode ser acomodada pelo multiplexador substituindo alguns bits de justificação por bits de informação. Em contrapartida, qualquer redução do débito do canal de entrada pode ser compensada adicionando mais bits de justificação.

Outra opção consiste em usar **justificação positiva/negativa.** Com esta técnica, o relógio de leitura opera à velocidade nominal do canal de chegada, de modo que quando o débito do tributário ultrapassa o seu valor nominal, há bits de informação que são removidos, e transmitidos em posições predefinidas na trama, como bits de justificação negativa. Por sua vez, quando o débito de entrada decresce são enviados bits de justificação, exactamente como no caso da justificação positiva. Os sistemas plesiócronos europeus usam unicamente justificação positiva, embora o ITU-T tenha também normalizado sistemas plesiócronos com justificação positiva/negativa.

Na Figura 5.10 apresenta-se o esquema de blocos de um multiplexador TDM, com capacidade para multiplexar *N* tributários. As sequências digitais provenientes dos diferentes tributários passam por uma memória elástica, de modo a serem sincronizadas com a frequência de relógio do multiplexador, e servem em seguida de entrada a um multiplexador síncrono (MUX). O multiplexador deve também executar as funções de justificação referidas anteriormente, assim como de sincronização dos diferentes canais. Devido a estas funções será de esperar que o débito binário de saída D_0 seja superior ao somatório dos débitos binários dos tributários de entrada.

5.2.2 Estrutura das tramas das segundas hierarquias plesiócronas

A trama de um sinal E2 é constituída por 848 dígitos binários (bits). Esta trama pode-se considerar estruturada em quatro sectores (sub-tramas) S_1, S_2, S_3, S_4, compreendendo cada um 212 bits. A repartição dos bits no sector S_1 é a seguinte:

- Os dez primeiros bits constituem o padrão de enquadramento de trama (1111010000);

- Os dois bits seguintes (nº 11 e 12) são bits de reserva (o bit nº 11 é usado para a transmissão de um alarme distante);

- Os 200 bits seguintes são destinados à transmissão de informação; eles constituem 50 grupos consecutivos idênticos, de quatro bits cada; como se usa interposição de bit em cada grupo, tem-se um elemento binário por tributário, ou seja o bit 1 é atribuído ao tributário 1, o bit 2 ao tributário 2 e assim sucessivamente.

Os sectores S_2, S_3 e S_4 são idênticos. Cada um destes sectores contém 212 bits fraccionados em 53 grupos de 4 bits cada. O primeiro grupo de cada sector corresponde aos bits de indicação de justificação. Cada bit deste grupo é alocado a um tributário, ou seja o bit 1 ao tributário 1 e assim sucessivamente. Os grupos seguintes são grupos de informação e têm a mesma estrutura que os grupos de informação do sector 1.

Os bits de justificação são introduzidos, quando é necessário, no primeiro grupo de informação do sector S_4, e ocupam a posição do bit de informação do tributário correspondente. A emissão de um pedido de justificação é assinalado colocando os três

bits indicadores de justificação correspondentes no estado lógico "um". Por sua vez, na mesma trama o bit de justificação ocupa o lugar destinado ao primeiro bit de informação do sector S_4 do tributário justificado.

A Tabela 5.5 mostra a estrutura da trama E2 descrita. Os bits F_i correspondem aos bits do padrão de enquadramento (com valor lógico i), o bit X é um bit de alarme, o bit Y é um bit de reserva, os bits I_i são bits de informação dos tributários, os bits C_{ij} são bits de indicação de justificação do canal i, e finalmente J_i são bits de justificação do canal i. Como exemplo, note-se que, a presença de um bit de justificação no canal 2 é indicada fazendo C_{21}=1, C_{22}=1 e C_{23}=1.

Tabela 5.5 Estrutura da trama E2 (8.448 Mb/s)

F_1	F_1	F_1	F_1	F_0	F_1	F_0	F_0	F_0	F_0	X	Y	I_{13}	....	I_{212}
C_{11}	C_{21}	C_{31}	C_{41}	I_5		....								I_{212}
C_{12}	C_{22}	C_{32}	C_{42}	I_5		....								I_{212}
C_{13}	C_{23}	C_{33}	C_{43}	J_1	J_2	J_3	J_4	I_9	....					I_{212}

A periodicidade com que os bits vazios podem ser inseridos nos *time-slot* predestinados define uma **trama de justificação** de duração T_j (não necessariamente igual ao período de amostragem T_a). O débito máximo de justificação para um determinado tributário é dado por

$$D_j = \frac{1}{T_j} \tag{5.3}$$

Assim, no caso do E2 tem-se uma trama de justificação com T_j ≈100.4 μs, portanto diferente do período de amostragem T_a =125 μs. Estes resultados são consequência do facto de a multiplexagem de segunda ordem ignorar a estrutura das tramas primárias, considerando simplesmente o fluxo dos bits. Para a mesma trama tem-se D_j ≈ 9962 bits/s.

Tabela 5.6 Estrutura de trama DS2 (6.312Mb/s)

M_0	I_{1-48}	C_{11}	I_{1-48}	F_0	I_{1-48}	C_{12}	I_{1-48}	C_{13}	I_{1-48}	F_1	J_1	I_{2-48}			
M_1	I_{1-48}	C_{21}	I_{1-48}	F_0	I_{1-48}	C_{22}	I_{1-48}	C_{23}	I_{1-48}	F_1	I_1	J_2	I_{3-48}		
M_1	I_{1-48}	C_{31}	I_{1-48}	F_0	I_{1-48}	C_{32}	I_{1-48}	C_{33}	I_{1-48}	F_1	I_1	I_2	J_3	I_{4-48}	
X	I_{1-48}	C_{41}	I_{1-48}	F_0	I_{1-48}	C_{42}	I_{1-48}	C_{43}	I_{1-48}	F_1	I_1	I_2	I_3	J_4	I_{5-48}

A Tabela 5.6 descreve a estrutura de uma trama DS2, a qual é constituída por 1176 bits. Como se vê, a estrutura desta trama é significativamente diferente da E2. Em primeiro lugar, tal como no DS1, o enquadramento da trama é distribuído. Além disso, têm-se dois níveis de enquadramento. O nível principal é estabelecido pelo padrão obtido

alternando F_0, F_1, F_0….O desmultiplexador usa este padrão para sincronizar a trama. Depois de ter obtido este sincronismo, procura o padrão estabelecido pelos bits M_i, (com valor lógico i). Este último nível, que divide a trama em quatro sub-tramas, é necessário para permitir identificar os bits C_{ij} e J_i, que têm a mesma definição que no caso anterior. O quarto bit M, designado por bit X, não é usado para fins de enquadramento, mas sim como elemento de serviço (alarme).

Exemplo 5.2

Um valor típico para a precisão dos relógios dos tributários da primeira hierarquia DS1 é de 50 ppm (partes por milhão). Pretende-se demonstrar que a presença de 1 bit de justificação por tributário na trama DS2 é suficiente para compensar as flutuações dos débitos associados à precisão referida, tendo presente que o débito nominal do DS1 é igual a 1.544 Mbit/s.

Solução: Cada trama DS2 contém 288 bits de informação para cada um dos canais (tributários), os quais se reduzem para 287 quando se usa justificação. Assim, o débito binário máximo por canal é igual a

$$\frac{288 \times 6.312\text{Mbit/s}}{1176} = 1.5458\,\text{Mbit/s}$$

enquanto o débito binário mínimo vem dado por

$$\frac{287 \times 6.312\text{Mbit/s}}{1176} = 1.5404\,\text{Mbit/s}$$

Esses débitos correspondem a uma flutuação relativamente ao débito binário nominal (1.544 Mbit/s), respectivamente de 1.8 kbit/s e – 3.6 kbit/s. Ou seja, 1 bit de justificação tem capacidade para acomodar flutuações dentro desses limites. Por sua vez, um relógio com uma tolerância (estabilidade) de 50 ppm, irá originar flutuações no débito binário de $\frac{50}{10^6} \times 1.544\text{Mbit/s} = 77.2\,\text{bit/s}$, as quais estão perfeitamente integradas nos limites permitidos por 1 bit de justificação.

5.2.3 Indicação de justificação

Como já se viu, cada *time-slot* destinado a justificação tem associado um conjunto de bits indicadores de justificação (bits C), que permitem ao desmultipexador interpretar adequadamente a informação contida nesse *time-slot*. Se um bit de informação é considerado erradamente bit de justificação, ou vice-versa, todos os bits da trama primária do tributário em questão estarão errados, já que a detecção incorrecta conduz à perda de bits de informação, ou então, à introdução de bits estranhos à mensagem.

Para proteger o sistema contra possíveis erros de transmissão usa-se normalmente um padrão de três bits para indicar a presença ou ausência de justificação, sendo usada na recepção uma descodificação baseada na decisão por maioria. Isto significa que a maioria dos C bits, neste caso dois em três, determina se o padrão transmitido foi 000, ou 111.

Admitindo que p é a probabilidade média de um bit individual ser detectado incorrectamente, e que os erros são independentes, tem-se que a probabilidade de tomar uma decisão errada sobre o padrão indicador de justificação (PIJ) é dada por

$$P_e(PIJ) = \sum_{i=(n+1)/2}^{n} \binom{n}{i} p^i (1-p)^{n-i} \tag{5.4}$$

onde n é o número de bits C que constituem o padrão indicador de justificação. Para os casos de interesse prático, em que a probabilidade de erro p é muito reduzida, a expressão anterior pode ser aproximada por

$$P_e(PIJ) \approx \binom{n}{x} p^x \tag{5.5}$$

onde $x=(n+1)/2$.

Para o caso em que se usa um padrão de três bits, a presença de erros no canal de transmissão pode levar à recepção das seguintes sequências: 000, 001, 010, 100, 011, 110, 101, 111. Como se sabe, destas sequências só duas correspondem a padrões correctos: 000 e 111. Assim, de acordo com a regra de decisão por maioria, sempre que o número de bits iguais a zero é superior ou igual a 2, admite-se que não há bit de justificação e decide-se pela sua presença em caso contrário. Deste modo, para as primeiras quatro sequências não há remoção do bit de justificação, enquanto para as quatro últimas esse bit é eliminado.

De acordo com a regra de decisão referida, apenas se toma uma decisão errada sobre o PIJ, quando ocorrem dois ou três erros no padrão de indicação de justificação. Assim, usando (5.4) tem-se

$$P_e(PIJ) = \binom{3}{2}(1-p)p^2 + p^3 \tag{5.4}$$

que nos casos em que p é muito pequeno se reduz a $P_e(PIJ) \approx 3p^2$.

Exemplo 5.3

Considere que usa uma segunda hierarquia europeia e que o canal de transmissão conduz a uma probabilidade de erro de 10^{-5}. Para cada um dos tributários E1 determine o tempo médio entre erros resultantes de erros no PIJ.

Solução: Usando a formulação anterior e tendo presente que o padrão de indicação de justificação é constituído por três bits, conclui-se que, $P_e(PIJ) \approx 3x10^{-10}$, ou seja, no canal E1 directamente afectado em média 3 tramas em cada 10^{10} estão erradas. Se se atender que a duração de uma trama E2 é de 100.4 μs então o tempo médio entre erros é de aproximadamente 93 h.

5.2.4 Perda e aquisição de enquadramento

Como já se referiu, o padrão de enquadramento de trama (PET) permite ao desmultiplexador identificar com precisão o início e o fim de uma trama, e deste modo definir com rigor as palavras de 8 bits correspondentes a cada canal. A função de enquadramento de trama é também referida como **sincronização de trama.**

O padrão de enquadramento é constituído por uma sequência binária, que apresenta uma fraca probabilidade de ocorrer durante a trama. A configuração do padrão e o seu comprimento variam em função da hierarquia e da norma de referência (ver Tabela 5.7). O comprimento da sequência correspondente ao PET deve resultar de um compromisso entre os seguintes factores:

- Um padrão de enquadramento longo tem maior probabilidade de ser recebido errado (desvantagem);

- Um padrão de enquadramento longo conduz a uma eficiência de transmissão mais baixa, pois aumenta o tempo necessário para transmitir o sinal (desvantagem);

- Um padrão de enquadramento longo torna mais eficiente a sincronização de bit, quando esta é feita à custa do mesmo padrão que a sincronização de trama (vantagem);

- Um padrão de enquadramento longo tem uma menor probabilidade de que os dados transmitidos coincidam com a sequência de enquadramento (vantagem).

Tabela 5.7 Recomendações da ITU-T para enquadramento de trama para a hierarquia europeia

Nível Hierárquico	Recomendação do ITU-T	Número de bits entre PET	Padrão de enquadramento	Número de violações para mudar de estado	
				K	*M*
E1	G 732	512	0011011 (7 bits)	3 ou 4	3
E2	G 742	848	1111010000 (10 bits)	5	2
E3	G 751	1536	1111010000 (10 bits)	4	3
E4	G.751	2928	111110100000 (12 bits)	4	3

A detecção do padrão de enquadramento pode ser efectuada usando um detector do tipo apresentado na Figura 5.11. O comparador tem por função comparar as sequências binárias transmitidas com o padrão de enquadramento armazenado numa memória. Por cada acordo entre 1 bit da sequência e um bit do padrão armazenado o comparador gera um 1, o qual vai incrementar de uma unidade um contador. Se o número de contagens é igual ao comprimento do PET, o padrão foi localizado.

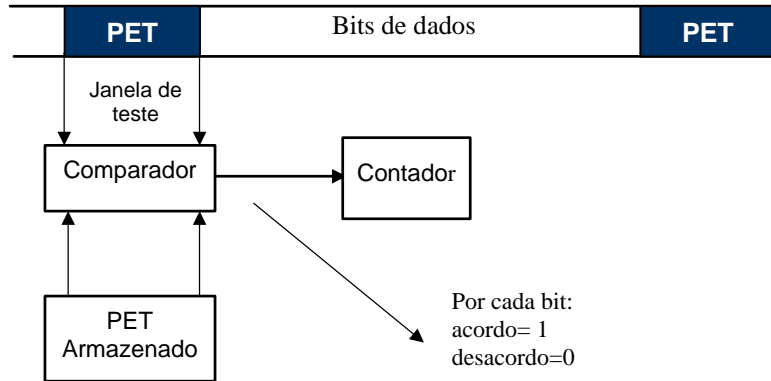



Figura 5.11 Configuração básica de um detector de padrões de enquadramento de trama.

Um circuito de sincronismo de trama tem dois estados básicos de operação, como se mostra na Figura 5.12. Quando o PET é observado continuamente na posição esperada diz-se que o circuito está em **estado de manutenção**. Quando ocorrem K violações sucessivas do PET o circuito muda para o **estado de procura.** A presença do circuito no estado de procura, significa que o sistema TDM em questão perdeu o sincronismo de trama, e por isso, durante todo o período em que estiver nesse estado esse sistema é colocado fora de serviço. No estado de procura o circuito pesquisa sistematicamente a sequência binária transmitida no sentido de encontrar o PET. Quando este padrão é encontrado em M tramas sucessivas, o circuito volta de novo para o estado de manutenção.

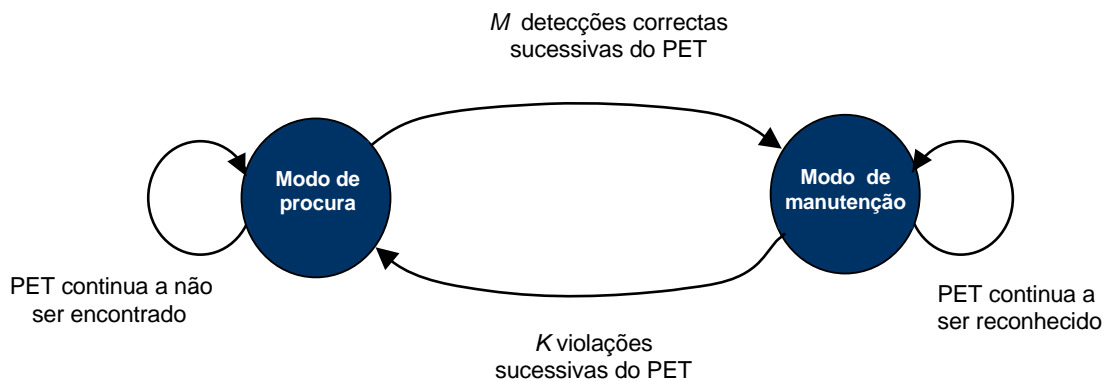



Figura 5.12 Diagrama de transição de estados de um circuito de sincronismo de trama.

A eficiência de uma determinada estratégia de sincronização de trama pode ser caracterizada especificando os seguintes parâmetros:

- Tempo em que uma trama está sincronizada, ou seja, o tempo médio entre declarações de perda de enquadramento;

- Tempo necessário para declarar a perda de enquadramento;

- Tempo necessário para readquirir o sincronismo, ou seja, tempo médio de aquisição de enquadramento;

Para quantificar estes parâmetros considere-se uma trama de comprimento igual a *L* bits, dos quais *N* correspondem ao padrão de enquadramento. Seja de novo a probabiliade de detectar um bit erradamente *p*. Assume-se que os erros de bit são devidos a ruído encontrado na transmissão, e que estes ocorrem aleatoriamente. Para além disso, considera-se, também, que um erro em qualquer um dos *N* bits é suficiente para errar o padrão de enquadramento. Nestas circunstâncias, a probabilidade de erro no padrão de enquadramento, P_p ,é dada por

$$\begin{aligned} P_p &= 1 - \Pr\{\text{padrão recebido sem erros}) \\ &= 1 - (1 - p)^N \end{aligned} \tag{5.7}$$

Para $p \ll 1$, esta expressão é bem aproximada por $P_p \approx Np$ (que contabiliza apenas a ocorrência de erros simples). A probabilidade de declarar perda de enquadramento, P_E, é igual à probabilidade de *K* padrões serem recebidos com erro, o que conduz a

$$P_E = (P_p)^K = \left[1 - (1 - p)^N\right]^K \approx (Np)^K \tag{5.8}$$

Se são transmitidas *F* tramas por segundo, então o tempo médio entre a declaração de perdas de enquadramento é dado por

$$T_E = \frac{K}{FP_E} \approx \frac{K}{F(Np)^K} \tag{5.9}$$

verificando-se, por conseguinte, que quanto maior é a sequência de enquadramento (*N*), menor será o tempo entre perdas de enquadramento, tal como se referiu anteriormente.

Vai-se em seguida calcular o tempo necessário para declarar a perda de enquadramento. Quando o enquadramento é perdido (devido à existência de *K* padrões errados), o sistema receptor pode não noticiar esta perda, pois os dados transmitidos são aleatórios e há uma probabilidade 2^{-N}, que uma sequência de *N* dígitos binários imite o PET. A probabilidade de detectar a perda de enquadramento será, assim, dada por

$$P_d = \left(1 - 2^{-N}\right)^K = 1 - K2^{-N} + ... \tag{5.10}$$

Interessa que a probabilidade P_d, esteja muito próxima de 1, o que implica necessariamente que,

$$K.2^{-N} \ll 1$$

Se a condição anterior é verificada, ou seja, se a probabilidade de encontrar uma sequência binária que imita o padrão de enquadramento é desprezável, o tempo necessário

para declarar a perda de enquadramento, T_d, é aproximadamente igual ao tempo necessário para transmitir *K* tramas.

$$T_d = \frac{K}{F(1-2^{-N})^K} \approx \frac{K}{F} \qquad (5.11)$$

Outro parâmetro importante é o tempo médio de aquisição de enquadramento de trama T_a, ou seja, o tempo necessário para procurar ao longo da sequência de dados (*N* bits de cada vez) até que o PET é encontrado e depois confirmado. Para diminuir a probabilidade de falso enquadramento, resultante do possível aparecimento de uma sequência de dados igual à de enquadramento, apenas se considera que o enquadramento foi adquirido, quando se encontra o PET em *M* tramas consecutivas (confirmação).

Para calcular T_a, considera-se o pior caso, que corresponde à situação em que se inicia a pesquisa no bit consecutivo ao primeiro bit do verdadeiro padrão de trama. Tem-se, assim, um número mínimo de M tramas para adquirir o enquadramento. Porém, este procedimento complica-se, porque existe sempre a probabilidade 2^{-N}, de aparecer uma sequência binária que imite o PET. Quando isto ocorre, a procura é suspensa até à trama seguinte, no sentido de confirmar o padrão encontrado. Normalmente, basta um intervalo de trama para detectar um falso padrão. Considera-se que em média, se têm *h* suspensões. Neste caso

$$T_a = \frac{M+h}{F} \qquad (5.12)$$

Para calcular *h*, deve-se ter presente, que no pior caso devem examinar-se *L*+*h* padrões, sendo *h*, de entre estes, detectados como padrões de enquadramento, de modo que

$$\frac{h}{L+h} = 2^{-N} \qquad (5.13)$$

o que permite escrever

$$h = \frac{L}{2^N - 1} \qquad (5.14)$$

Note-se que *h* diminui com *N*, de acordo com o afirmado anteriormente. O tempo médio total (pior caso) para detecção de perda e aquisição de enquadramento será dado por

$$T_t = T_d + T_a = \frac{K + M + \dfrac{L}{2^N - 1}}{F} \qquad (5.15)$$

5.3 Hierarquia Digital Síncrona

Com a introdução da tecnologia PCM e da hierarquia digital plesiócrona em meados dos anos sessenta iniciou-se a evolução das redes de telecomunicações no sentido da sua digitalização. Os primeiros sistemas PDH correspondiam a ilhas digitais num mundo analógico, com a desvantagem destas ilhas não terem nenhuma ligação (sincronismo) entre si, não podendo, por isso, derivar o relógio dos multiplexadores do relógio da rede. Esta limitação da tecnologia PDH, que está directamente associada às condições subjacentes à sua génese, dificultou a sua aplicação no novo mundo das telecomunicações, dominado pelo paradigma da digitalização total.

Embora fosse teoricamente possível ultrapassar as limitações dessa tecnologia através do projecto de uma nova geração PDH, a tendência geral foi no sentido da concepção de um novo sistema com uma filosofia diferente, ou seja, a Hierarquia Digital Síncrona ou SDH (*Synchronous Digital Hierarchy)*. Essa nova hierarquia foi definida pela ITU-T em 1990. A sua normalização nos EUA foi anterior e tem a designação de SONET (*Synchronous Optical NETwork*), pela simples razão de que nesta rede as interfaces são normalizadas para sinais ópticos. A filosofia associada à SDH conduz à construção de uma rede em que todos os elementos operam de modo síncrono sobre o controlo de um relógio central da rede. A nova hierarquia apresenta um conjunto de vantagens, que são analisadas em seguida por oposição às limitações da PDH:

- **Débitos de transmissão elevados:** Na hierarquia plesiócrona não havia normas acima dos 140 Mbit/s e havia incompatibilidade entra as normas europeias, americanas e japonesas. Com a SDH a normalização é estendida até 10 Gbit/s e a maior parte das incompatibilidades referidas são ultrapassadas.
- F**lexibilidade :** Como se viu, na multiplexagem plesiócrona é necessário introduzir bits de justificação e bits de controlo de justificação de modo a ultrapassar os problemas associados às ligeiras diferenças entre as frequências dos diferentes relógios. Este facto, associado à utilização de multiplexagem por interposição de bit, impedem o conhecimento dos limites dos diferentes canais e tornam o sistema muito inflexível. Uma operação, conceptualmente de grande simplicidade, como seja, por exemplo, a extração, ou adição de um canal a 2 Mbit/s de uma trama E4, requer a concatenação de uma grande quantidade de multiplexadores/desmultiplexadores (ver Figura 5.13), tornando a operação dispendiosa e economicamente inviável. No caso da SDH todos os relógios estão sincronizados com um relógio mestre central. Como consequência, não é necessário introduzir bits de justificação e os débitos binários do sinal multiplex são múltiplos inteiros do débito básico. Estes aspectos, associados ao facto de a multiplexagem ser realizada por interposição de octeto, permitem localizar com facilidade a posição dos diferentes canais na trama e extrair (ou inserir), com relativa simplicidade, tributários de débitos reduzidos. Como consequência, os multiplexadores de inserção/extracção são de fácil construção, apresentando-se como um dos elementos de rede mais importantes da tecnologia SDH.

- **Interligação:** Uma limitação significativa da PDH reside no facto de não especificar nenhum formato padrão para a transmissão dos sinais multiplexer nas vias de transmissão ópticas. Como consequência, diferentes fabricantes usavam diferentes

códigos de linha e diferentes interfaces ópticas para os seus produtos, tornando praticamente impossível a tarefa de interligar por vias ópticas equipamento de diferentes fabricantes. As interfaces ópticas da SDH foram fruto de uma profunda normalização (códigos, tipo de fibra, nível de potência, características das fontes, etc), facilitando, assim, a tarefa de interligação.

- **Gestão de rede e monitorização de desempenho:** A SDH possui um grande número de octetos para funções de gestão de rede, monitorização de desempenho, identificação do tipo de tráfego transportado e identificação e reportagem de falhas na rede. Há, especificamente, um conjunto de octetos que constituem um canal de comunicação de dados para transmitir informação de gestão entre os diferentes elementos de rede, que permite que estes sejam controlados e monitorizados por um sistema de gestão central através de uma rede de gestão de telecomunicações designada normalmente por TMN (*Telecommunications Network Management*). Na PDH as funcionalidades descritas são praticamente inexistentes.

- **Fiabilidade:** As quantidades de informação excepcionalmente elevadas transportadas pelas redes de telecomunicações modernas, fazem com que os padrões de exigência relativos à fiabilidade da rede sejam muito elevados. A topologia fisica (normalmente anel) e as técnicas de protecção específicas da SDH, permitem garantir esses níveis de exigência, assegurando tempos de restauro de serviço em presença de falhas da ordem dos 50 ms, valores muito inferiores aos típicos das redes PDH, que variam entre vários segundos até várias minutos.

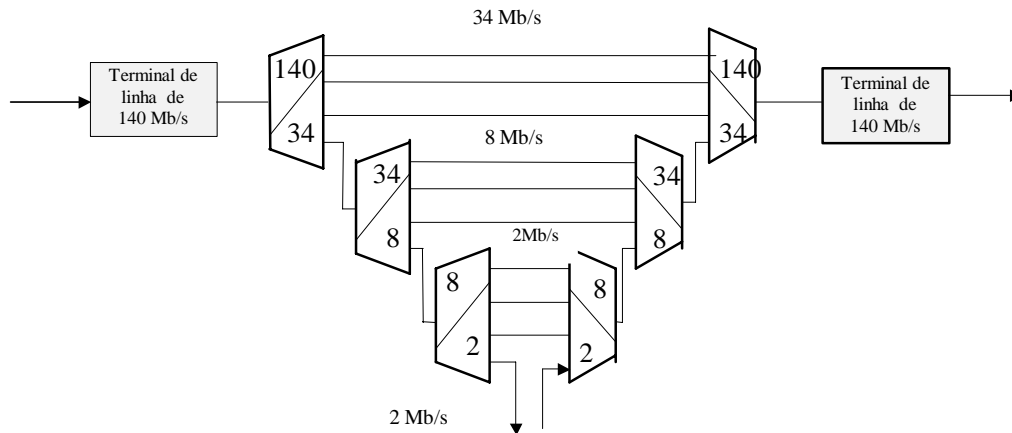



Figura 5.13 Extracção de um canal a 2 Mb/s de uma quarta hierarquia plesiócrona.

O sinal SDH básico designa-se por módulo de transporte síncrono de nível 1, ou STM-1 (*Synchronous Transport Module level 1*), ao qual corresponde um débito binário de 155.52 Mb/s. Os débitos binários das hierarquias superiores são múltiplos deste valor, por um factor de $N=4^n$ (n=1, 2, 3, 4) conduzindo aos sinais STM-*N* como se indica na Figura 5.14.[63]

[63] Note-se que nessa figura os débitos estão aproximados.

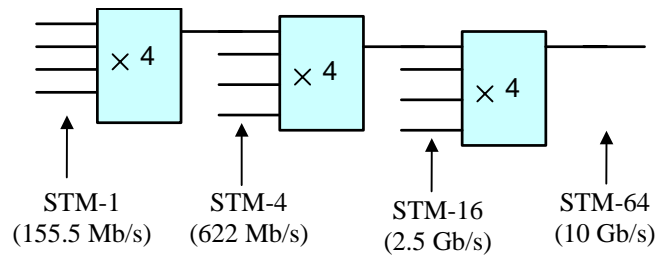
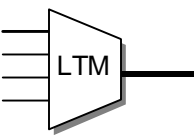
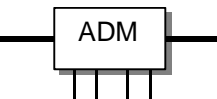
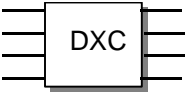
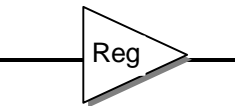



Figura. 5.14 Hierarquia digital síncrona.

No caso do SONET o sinal básico designa-se por STS-1 (*Synchronous Transport Signal level 1*) e corresponde a 51.84 Mb/s, concluindo-se assim que um sinal STS-3 do SONET é equivalente ao sinal STM-1 do SDH (ver Tabela 5.7). Note-se que, o sinal STS-*N* é um sinal eléctrico e particularmente para débitos elevados só existe no interior do equipamento. A interface para outro tipo de equipamento é obtida baralhando (*scrambling*) o sinal STS-*N* e convertendo-o para o domínio óptico. O sinal correspondente, de acordo com a nomenclatura do SONET, designa-se por OC-*N* (*optical carrier level N*). Assim, por exemplo, ao STS-3 corresponde no domínio óptico o sinal OC-3.

O débito do sinal básico do SONET foi escolhido de modo a acomodar os sinais plesiócronos DS1 e DS3, enquanto o débito básico da SDH foi escolhido de modo a acomodar os sinais plesiócronos E1, E3 e E4, assim como a hierarquia americana (com excepção do DS4). Os débitos dos sinais do SONET de hierarquias superiores à primeira coincidem com os débitos da SDH.

Tabela 5.8 Características dos sinais síncronos mais comuns.

Sinal SDH	Sinal SONET	Débito (Mb/s)
STM-0*	STS-1	51.84
STM-1	STS-3	155.52
STM-4	STS-12	622.08
STM-16	STS-48	2488.32
STM-64	STS-192	9953.28
* não está normalizado		

5.3.1 A SDH como Tecnologia de Transporte

Na Secção 1.4.2 do Capítulo 1 referiu-se que é usual identificar numa rede de telecomunicações uma camada de rede de serviço e uma camada de rede de transporte, e que camada de rede de transporte compreende as funcionalidades necessárias para proporcionar transmissão, multiplexagem, encaminhamento, protecção, supervisão, e aprovisionamento de capacidade dos sinais provenientes das camadas de serviço. A SDH é uma tecnologia que permite implementar todas essas funcionalidades e por isso é considerada como uma tecnologia de transporte.

5.3.2 Elementos da Rede de Transporte SDH

Uma rede de transporte SDH é constituída por vários elementos de rede (nós) interligados entre si, normalmente por vias de transmissão ópticas, segundo uma certa topologia física, e ligados a um sistema de gestão de rede pelo canal de comunicação de dados. Os elementos de rede SDH fundamentais são os multiplexadores, comutadores de cruzamento (*cross-connects*) e regeneradores. Os multiplexadores podem ainda operar em modo terminal ou em modo de inserção/extracção (*Add/Drop*). As características desses elementos e as suas principais funções estão esquematizadas na Tabela 5.9.

Tabela 5.9 Definição dos elementos da rede SDH.

Tipo	Figura	Função
Multiplexadores Terminais de Linha LTM (*Line Terminal Multiplexer*)	LTM	É usado para combinar tributários PDH, ou mesmo SDH de modo a gerar um sinal STM-*N* de débito mais elevado. Estes elementos são a principal adaptação entre hierarquia PDH e a hierarquia SDH.
Multiplexadores de Inserção/Extracção ADM (*Add/Drop Multiplexer*)	ADM	Permite extrair ou inserir sinais plesiócronos ou síncronos de baixo débito no sinal STM-*N* que passa, e não termina no equipamento. Este elemento de rede representa uma das grandes vantagens da nova hierarquia: um sinal de um nível hierárquico superior não necessita de ser desmultiplexado para que seja efectuada uma extracção ou adição de um tributário.
Comutador de Cruzamento Digital DXC (*Digital Cross Connect System*)	DXC	Permitem estabelecer ou modificar as ligações entre vários canais de entrada e de saída, actuando desde o nível E1 até ao nível STM-1. Note-se que a sua reconfiguração é realizada por controlo do sistema de gestão, enquanto os comutadores respondem à sinalização do cliente.
Regenerador SDH REG (*Regenerator*)	Reg	Para distâncias de transmissão superiores a 60 km é necessário recorrer a regeneradores. Estes, para além de fazerem a simples regeneração do sinal, têm capacidades de monitorização de desempenho e capacidades de alarme e vigilância.

5.3.3 Arquitecturas de Redes SDH

As redes SDH são construídas interligando diferentes elementos de rede com vias de transmissão ópticas, segundo uma determinada topologia física. As topologias físicas mais usadas são a cadeia, o anel e a malha, as quais irão ser analisadas em seguida.

5.3.3.1 Rede em Cadeia

Uma rede em cadeia é definida como um conjunto de nós interligados de modo a formar uma arquitectura linear. Esta estrutura é normalmente implementada usando multiplexadores terminais nos nós extremos e multiplexadores de inserção/extracção nos nós intermédios. A capacidade de cada secção da cadeia é igual à capacidade da cadeia (por exemplo STM-1, STM-4, STM-16 ou STM-64). A principal limitação desta topologia é sua reduzida conectividade, o que limita consideravelmente o número de técnicas de protecção que podem ser empregues.

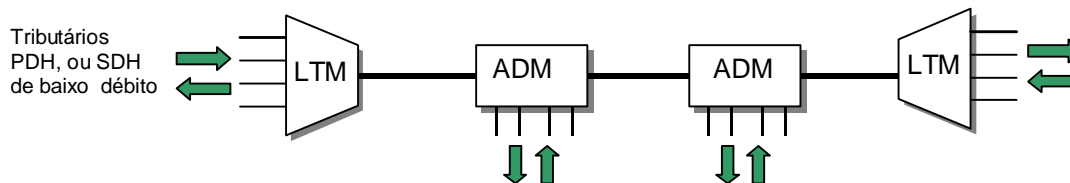



Figura 5.15 Rede SDH em cadeia

5.3.3.2 Rede em anel

A topologia em anel é a topologia mais usada na concepção de redes SDH. É definida como um conjunto de nós interligados de modo a formar uma malha fechada, sendo cada nó ligado a dois nós adjacentes, e é caracterizada por cada secção do anel ter a mesma capacidade de todo o anel, de modo idêntico à topologia em cadeia.

Os elementos de rede usados são essencialmente os multiplexadores de inserção/extracção. Estes componentes devem permitir um acesso completo a todo o tráfego de linha que circula no anel e estão equipados com uma matriz de comutação sem bloqueio de modo a permitir reconfigurar o elemento. A sua principal aplicação consiste em inserir ou extrair tráfego dos tributários a partir do sinal de linha STM-*N*.

As redes SDH em anel podem dividir-se em duas categorias gerais: anéis unidireccionais e anéis bidireccionais. Nos anéis unidireccionais todo o tráfego (bidireccional) de serviço é transportada ao longo do anel num único sentido (por exemplo no sentido horário). Ou seja, o sinal de chegada a um nó viaja numa direcção e o correspondente sinal de partida do mesmo nó segue na mesma direcção, de modo que o conjunto dos dois sinais ocupa toda a circunferência do anel. Num anel bidireccional o sinal de chegada a um nó viaja numa direcção e o correspondente sinal de partida viaja em direcção oposta, mas através do mesmo arco (secção) do anel. A topologia em anel tem a vantagem de suportar um grande número de técnicas de protecção, como se irá analisar na secção 5.3.10. As redes de transporte SDH não usam na generalidade dos casos um único anel, mas sim arquitecturas multi-anel, baseadas na interligação de vários

anéis. De modo a aumentar a sobrevivência da rede em presença de falhas essa interligação é feita no mínimo através de dois nós. Os nós de interligação podem ser multiplexadores de inserção/extracção ou comutadores de cruzamento.

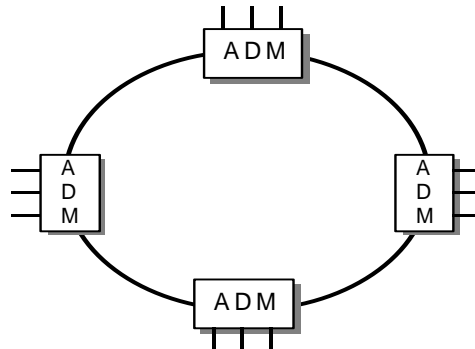



Figura 5.16 Rede SDH em anel baseada em ADMs.

5.3.3.3 Rede em malha

As redes de transporte SDH em malha, também designadas por redes emalhadas, são usadas sobretudo a nível do núcleo central da rede, ou por outras palavras, a nível da sua componente dorsal (*backbone*). Estas redes são construídas usando comutadores de cruzamento como nós de rede, interligados por vias de transmissão (vias ópticas + regeneradores).

Os nós DXC são geridos remotamente pelo sistema de gestão de rede, o que permite um aprovisionamento rápido de caminhos, aumentando, assim, a flexibilidade da rede. A gestão remota permite ainda que os nós DXC possam ser usados para implementar um sistema de restauro dinâmico, permitindo melhorar a disponibilidade da rede. Note-se que, o sistema de gestão intervém no processo de restauro da rede em presença de falhas, enquanto nos processos de protecção a rede recupera de modo autónomo, como se irá ver na secção 5.3.10.

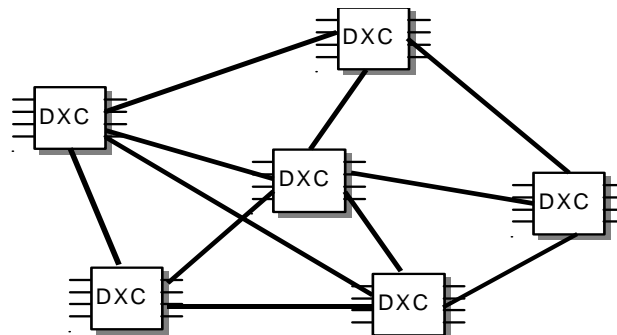



Figura 5.17 Rede em malha usando DXC.

5.3.4 Estrutura Estratificada da Rede

O fluxo de informação numa rede de transporte SDH é também baseado num modelo de camadas, constituído pela **camada de caminho** e pela **camada de transmissão**. A camada de caminho é responsável pelo transporte da informação entre os nós extremos da ligação. O cabeçalho de caminho (informação adicional introduzida pela camada de caminho necessária para realizar as funções requeridas para garantir a integridade da informação) é inserido no nó onde a ligação é originada e é retirado no nó destino. A camada de caminho ainda se subdivide em camada de caminho de ordem superior e camada de caminho de ordem inferior, dependendo da maior ou menor capacidade de transporte do caminho. A camada de transmissão subdivide-se, em camada de secção e camada física e a primeira subdivide-se, por sua vez, em secção de multiplexagem e secção de regeneração.

A camada de secção de multiplexagem multiplexa um certo número de caminhos numa única via entre dois nós. Por isso, esta camada é terminada em cada elemento de rede onde têm lugar operações de multiplexagem (LTM, ADM, DXC). A camada de secção de multiplexagem é também responsável por realizar certos tipos de comutação de protecção no caso de uma falha da via de transmissão ou de um nó, e ainda por funções de sincronização de relógio e manutenção. A camada de secção de regeneração é delimitada pelos regeneradores. É esta camada que é responsável pelo enquadramento das tramas e pela formação da trama básica da SDH e ainda pela monitorização dos regeneradores. A camada física é responsável pelo meio físico usado para transportar a informação, normalmente a fibra óptica, e pela adaptação do meio físico à camada de regeneração. Na Figura 5.18 exemplifica-se as terminações das camadas de caminho, secção de multiplexagem e regeneração para uma rede SDH em cadeia.

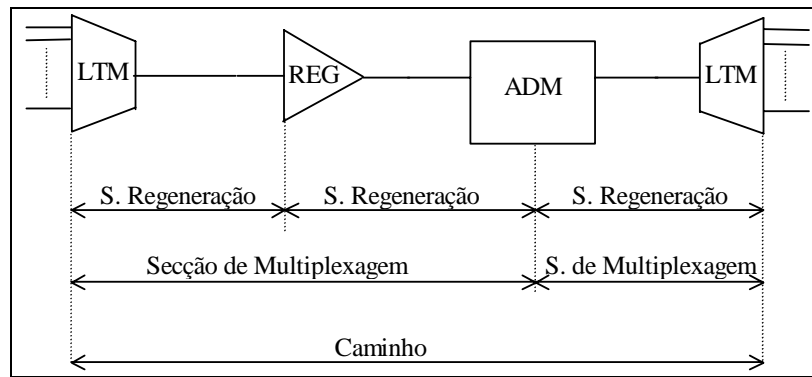



Figura 5.18 Definição das camadas de uma rede SDH em cadeia.

Cada camada terá de processar o seu próprio cabeçalho. Assim, no processo de formação da trama começa-se por adicionar aos dados recebidos da camada de rede de serviço o cabeçalho de caminho ou POH (*Path OverHead*), em seguida o cabeçalho de secção de multiplexagem ou MSOH (*Multiplexer Section OverHead*) e finalmente o cabeçalho da secção de regeneração ou RSOH (*Regenerator Section Overhead*) (ver Figura 5.19). Num regenerador só é processado o RSOH, ou seja, na entrada esse cabeçalho é retirado e é inserido na saída. Talvez seja útil fazer um breve comentário sobre o que se entende por

retirar o cabeçalho RSOH. Como se irá ver posteriormente, nesse cabeçalho vão um conjunto de canais a 64 kb/s, onde cada canal corresponde a um conjunto de 8 bits (octeto) transmitidos com uma periodicidade de 125 μs, que permitem à rede comunicar com o regenerador. Deste modo, retirar o cabeçalho RSOH significa que o regenerador passa a aceder directamente a esses canais, usando para seu proveito a informação que neles é transportada.

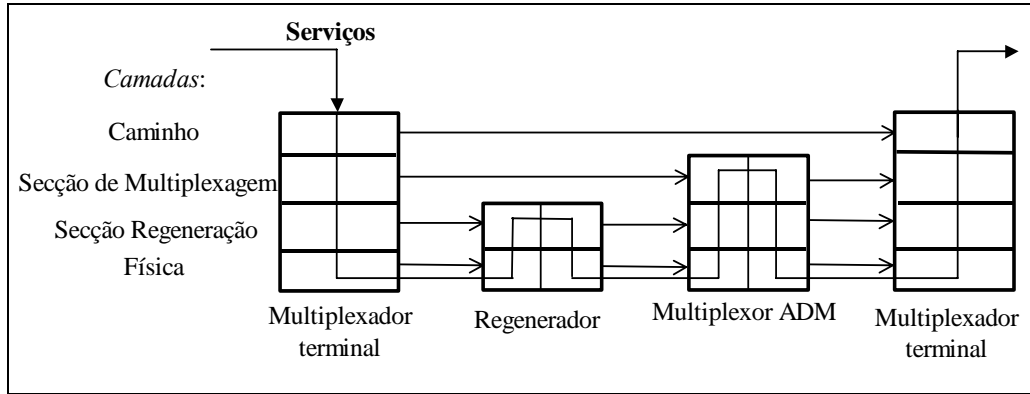



Figura 5.19 Ordem de processamento dos cabeçalhos.

Nos elementos de rede LTM, ADM e nos comutadores de cruzamento são processados os dois cabeçalhos de secção (RSOH e MSOH). Porém, no caso em que estes elementos de rede iniciam ou terminam um caminho, para além dos cabeçalhos de secção também é processado o cabeçalho de caminho. Também aqui, o termo processar tem o mesmo significado que anteriormente, ou seja, retirar o respectivo cabeçalho na entrada do elemento de rede e adicioná-lo na saída.

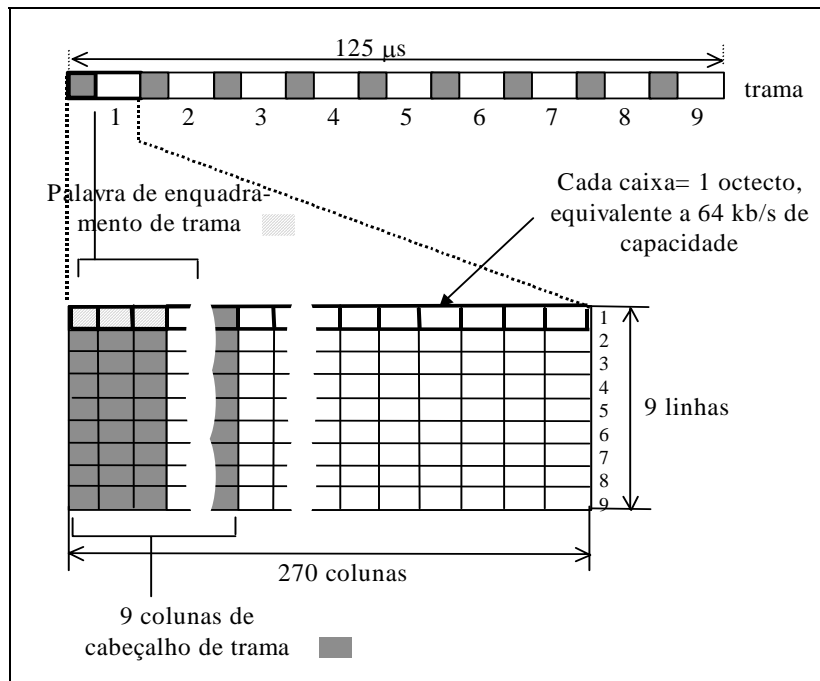



Figura 5.20 Estrutura da trama SDH (STM-1).

5.3.5 Estrutura da trama

A filosofia associada ao processamento de informação nas várias partes da rede SDH é completamente diferente da dos sistemas PDH. Assim, será de esperar que a estrutura das tramas também seja diferente. Aqui, a trama correspondente ao sinal STM-1, também designada por quadro básico, é usualmente representada numa forma bidimensional consistindo numa matriz de 9 linhas por 270 colunas (contendo 2430 octetos), como se representa na Figura 5.20. Este quadro é repetido cada 125 μs, o que implica um débito binário de 155.52 Mb/s, ou seja, um débito por octeto de 64 kb/s. A ordem de transmissão é a seguinte: começa-se pelos bits do canto superior esquerdo do quadro e transmite-se linha a linha até concluir a transmissão dos 2430 octetos.

Num quadro básico podem-se identificar três tipos de blocos: um bloco correspondente ao cabeçalho de secção, um bloco correspondente ao ponteiro e um terceiro bloco correspondente à carga digital transportada. As primeiras 9 colunas são reservadas para os primeiros dois blocos, enquanto as restantes 261 colunas ($261\times9\times64$ = 150.336 Mb/s) estão reservadas para o transporte da carga digital. Esta carga inclui a informação dos tributários plesiócronos, assim como octetos adicionais apropriados para organizar de modo adequado essa informação na trama STM-1.

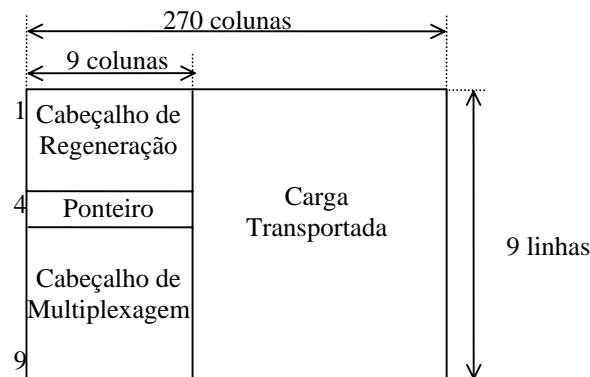



Figura 5.21 Estrutura da trama STM-1.

Como se ilustra na Figura 5.21, as primeiras 3 linhas do bloco correspondente ao cabeçalho de secção são usadas para o cabeçalho de secção de regeneração, e as cinco últimas para o cabeçalho de secção de multiplexagem. A quarta linha corresponde ao bloco do ponteiro. O ponteiro permite identificar o início da carga digital e acomodar justificações positivas ou negativas, como se irá ver posteriormente.

Os sinais multiplex SDH de ordem superior são obtidos por interposição de octeto de vários sinais STM-1. Por exemplo, um sinal STM-N resulta da multiplexagem de N sinais STM-1. A trama correspondente ao sinal STM-N tem também a duração de 125μs e a sua representação bidimensional inclui 270×N colunas, cada coluna com 9 linhas. Tendo presente a estrutura do sinal STM-1, conclui-se facilmente que os blocos correspondentes ao cabeçalho de secção e ao ponteiro do sinal STM-N ocupam 9×N colunas, enquanto o bloco associado à carga transportada ocupa 261×N colunas.

Na figura 5.23 apresenta-se com detalhe a estrutura do cabeçalho de secção, do ponteiro e também do cabeçalho de caminho de ordem superior de um sinal STM-1. No cabeçalho de secção de regeneração o conjunto dos octetos A1 e A2 (A1=11110110 e A2=00101000) constituem o padrão de enquadramento de trama, sendo usados para definir o início da trama STM-1. Os octetos A1 e A2 são os últimos octetos a serem inseridos no processo de formação da trama, a qual é completamente baralhada (*scrambled*) antes desses octetos serem inseridos, para evitar longas sequências de 0s e 1s. J0 é designado por traço da secção de regeneração e é usado para transmitir um identificador de modo que um receptor de secção possa verificar a integridade da sua ligação a um determinado emissor. B1 é usado para monitorizar erros de bit a nível da secção de regeneração. Esta função é implementada usando um código de paridade de bits entrelaçados, ou código BIP (*bit interleaved parity*), também designado por BIP-8 porque estão disponíveis 8 bits para o transmitir[64]. O código BIP-8 é calculado sobre todos os bits da trama STM-1 anterior e o resultado do cálculo é colocado no octeto B1 da trama corrente antes do baralhamento. D1, D2 e D3 constituem um canal de comunicação de dados ou DCC (*data communication channel*) a 192 kbit/s, que é usado para transportar informação de gestão de rede entre e para os regeneradores. E1 (não confundir com a hierarquia plésiocrona E1) proporciona um canal de comunicação de voz a 64 kb/s entre os regeneradores para ser usado pelo pessoal de manutenção, enquanto F1 se designa por canal de utilizador e pode ser usado para diferentes tipos de aplicações, como sejam transmissão de alarmes remotos ou estabelecimento de ligações temporárias para transmissão de dados, etc. Os octetos X são reservados para uso nacional, enquanto os octetos Δ transportam informação dependente do meio de transmissão (fibra, feixes hertzianos, satélites). Os campos deixados em branco correspondem a octetos cujas funções ainda não foram normalizadas.

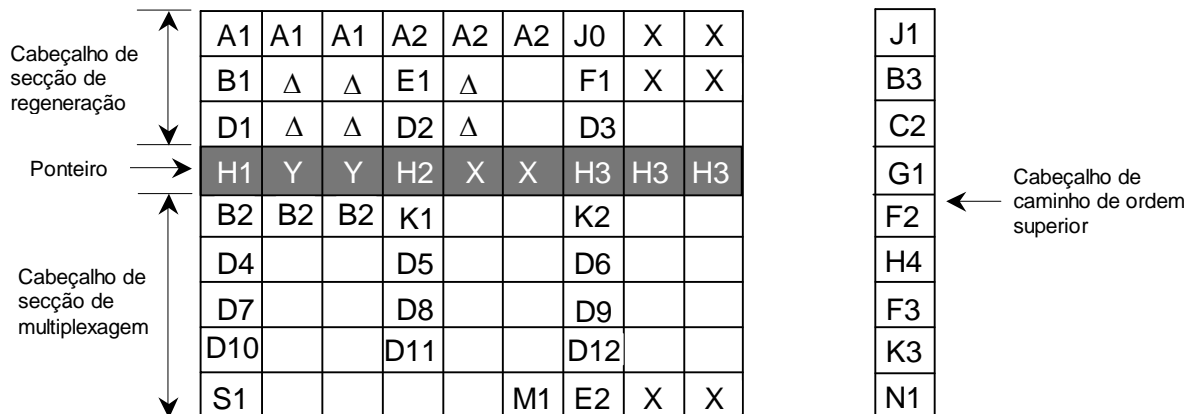



Figura 5.22 Estrutura dos cabeçalhos da trama STM-1.

Os octetos B2, D4-D12 e E2 do cabeçalho de secção de multiplexagem desempenham uma função análoga à dos octetos equivalentes do cabeçalho de secção de regeneração. Em particular, B2 é usado para transmitir um código BIP-24, que é calculado sobre todos os bits da trama anterior com excepção do cabeçalho de secção de regeneração, enquanto

[64] Consiste em adicionar em módulo 2, oito conjuntos de bits e colocar o resultado no octeco B1.

D4-D12 constituem um canal DCC a 576 kbit/s disponível para transmitir informação de gestão entre e para os elementos de rede que terminam a secção de multiplexagem. K1 e K2 são usados para a função de comutação de protecção automática, como se verá posteriormente, quanto se estudarem as diferentes técnicas de protecção. Por sua vez, o octeto S1 é usado para transmitir mensagens referentes ao tipo de relógio usado no processo de sincronização, enquanto o octeto M1 serve para transportar a informação referente ao número de erros detectados na secção de multiplexagem através dos octetos B2. O significado dos campos H1-H3 do ponteiro será explicado na secção 4.3.7.

Antes de explicar as funções dos diferentes octetos do cabeçalho de caminho é conveniente introduzir o conceito de contentor virtual, que tem um papel fulcral na estrutura de multiplexagem da SDH. O contentor virtual é obtido adicionando aos dados dos tributários plesiócronos o cabeçalho de caminho, isto depois de adaptar convenientemente os débitos desses dados aos requisitos da trama SDH, assunto que irá ser detalhado mais tarde. Relativamente ao cabeçalho de caminho, as funções dos diferentes octetos são as seguintes:

- **J1**: É o primeiro octeto do contentor virtual. A sua posição é indicada pelo ponteiro. É designado por traço do caminho e é usado para transmitir um identificador, de modo que a terminação de um determinado caminho possa verificar a integridade da sua ligação à fonte desse caminho.

- **B3**: Função de monitorização de erros. É usado para transmitir um código BIP-8, que é calculado sobre todos os bits do contentor virtual anterior e é transmitido no contentor virtual presente.

- **C2**: Indica o tipo de tráfego transportado pelo contentor virtual. (Ex: não equipado, tráfego plesiócrono E3, DS3 ou E4, ATM (*Asynchronous Transfer Mode*), FDDI (*Fiber Distributed Digital Interface*), HDLC (*High Level Data Link Control*), etc). Note-se que, para o transporte de tráfego IP sobre o SDH o sinal IP depois de sofrer algum processamento é mapeado numa trama HDLC.

- **G1**: Permite informar a fonte do estado e desempenho de um determinado caminho, usando a informação obtida pela terminação do caminho.

- **F2 e F3**: Reservados para o canal do utilizador, podendo ser acessível pelo utilizador terminal no estabelecimento de uma ligação.

- **H4**: Indicador de multitrama, no caso em que a carga digital é distribuída ao longo de várias trama, como é o caso do transporte dos contentores VC-2,VC-11 e VC-12.

- **K3**: Usado para comutação de protecção automática a nível de caminho.

- **N1**: Usado para monitorizar os caminhos, que se estendem por diferentes redes (ligações em cascata).

5.3.6 Estrutura de multiplexagem

A hierarquia digital síncrona foi concebida de modo a permitir agregar, quer o tráfego plesiócrono associado às hierarquias americanas e europeias, quer ainda o tráfego síncrono. Obteve-se, assim, uma nova hierarquia caracterizada pela sua grande flexibilidade, exigência primordial das redes de telecomunicações modernas. Em contrapartida, sacrificou-se a simplicidade estrutural das tramas plesiócronas, já que tanto as tramas SDH, como o processo da sua formação são bastante complexos. Para melhor compreender esse processo, o qual é representado na Figura 5.24, é usual definirem-se as seguintes sub-estruturas modulares:

- **Contentor de nível *n*** ou C-*n* (*Container*-*n*), com *n* de 1 a 4, é a unidade básica da estrutura SDH, sendo usada para transportar a informação dos tributários correspondentes às hierarquias plesiócronas europeias e americanas.

- **Contentor Virtual de nível n,** ou VC-n (*Virtual Container*-n), com *n* de 1 a 4, é a unidade fundamental para o transporte de informação a nível de caminho e consiste num contentor mais o seu cabeçalho de caminho, ou numa associação de diferentes grupos de unidades tributárias. Tendo presente a estrutura estratificada da camada de caminho, os VC-1 e VC-2 são considerados VC de ordem inferior, enquanto os VC-3 e VC-4 são considerados contentores de ordem superior. Nos países CEPT o VC-3 também é considerado como VC de ordem inferior.

- **Unidade Tributária de nível n,** ou TU-n (*Tributary Unit*-n), com n de 1 a 3, consiste num contentor virtual mais um ponteiro de unidade tributária. O ponteiro especifica o início do VC-n dentro do VC-n+1.

- **Unidade Administrativa de nível n,** ou AU-n (*Administrative Unit*-n), com n=3, 4, consiste num VC-n (n=3,4) mais um ponteiro de unidade administrativa. O ponteiro AU tem uma localização fixa no STM-1 e especifica o início do VC-n dentro do STM-1.

- **Grupo de Unidade Tributária de nível n,** ou TUG-n (*Tributary Unit Group*-n), com n=2, 3, resulta da combinação de várias unidades tributárias usando interposição de octeto, sem a adição de qualquer cabeçalho adicional.

- **Grupo de Unidade Administrativa**, ou AUG (*Administrative Unit Group*), resulta da combinação por interposição de octeto de várias AUs.

- **Módulo de transporte síncrono,** ou STM**,** é a subestrutura modular que é transmitida directamente sobre o meio físico. O STM-1 é obtido adicionando ao AUG o cabeçalho de secção. Em termos gerais o STM-N é obtido combinando N AUGs e N cabeçalhos se secção, usando interposição de octeto.

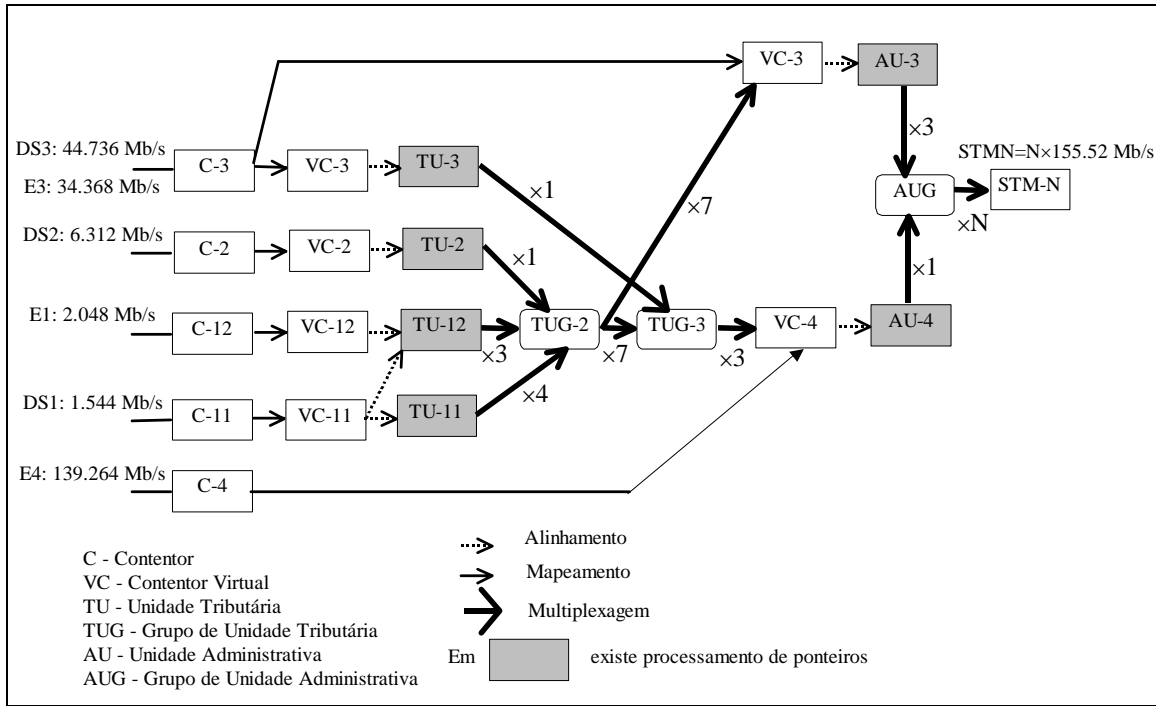



Figura 5.23 Estrutura de multiplexagem SDH.

Como se pode ver a partir da Figura 5.23 para transportar as hierarquias E1, E3 e E4 usam-se, respectivamente, os contentores C12, C3 e C4. Será de destacar que, o processo de inserção de um tributário plesiócrono num contentor envolve uma operação de justificação similar à que tem lugar nas multiplexagens plesiócronas de ordem superior. Contudo, neste caso, o canal plesiócrono é sincronizado não com o relógio do multiplexador, mas sim com a frequência do contentor, que por sua vez é síncrona com a estrutura básica da trama SDH. Para além dessa operação, é necessário adicionar octetos sem conteúdo informativo para adaptar o débito binário dos sinais PDH, com o débito binário dos contentores. Esta operação também é designada por alguns autores por justificação fixa.

5.3.7 O papel dos ponteiros das unidades administrativas

O contentor virtual é transportado intacto através da rede SDH desde o elemento de rede onde o caminho é originado (tráfego do tributário carregado) até ao elemento de rede onde o caminho é terminado (tráfego descarregado e entregue ao destinatário). Ao longo do percurso, os diferentes cabeçalhos (com excepção do cabeçalho de caminho) são extraídos e inseridos, sucessivamente. Ao formar-se uma nova trama, num determinado elemento de rede, todo o processo é sincronizado pelo relógio desse elemento de rede.

Como é pouco provável que esse relógio e o relógio do elemento de rede onde o contentor virtual foi formado estejam perfeitamente sincronizados em fase (numa rede síncrona há um sincronismo perfeito a nível de frequência), será de esperar a existência de ligeiras discrepâncias entre a fase da trama, e a fase do contentor virtual, as quais se vão acumulando ao longo da rede. A dessincronização de fase entre os relógios é devida a

variações no atraso de propagação, quando o caminho é estabelecido na rede de um único operador, ou ausência de sincronismo entre os relógios, quando o caminho é estabelecido entre redes de diferentes operadores, ou perdas de referência de relógio devido a falhas.

Para acomodar essas flutuações entre as fases as hierarquias síncronas foram concebidas assumindo que o contentor virtual não ocupa posições rígidas, mas que pode flutuar dentro da trama, o que implica que a localização do início do contentor virtual seja transmitida dentro da própria trama. Tendo como exemplo a AU-3, o início do contentor virtual VC-3, que é definido pelo octeto J1, é indicado pelo ponteiro da AU-3. O ponteiro da AU-3 é constuído pelos octetos H1, H2 e H3. A posição do octeto que se segue a H3 é numerada por 0. As diferentes posições são numeradas linha a linha até à posição 782. O octeto que ocupa esta posição, precede o octeto H1 da trama seguinte, já que como consequência lógica da filosofia usada, a transmissão de um contentor virtual ocupa sempre duas tramas. Na Figura 5.24 exemplifica-se essa situação tendo como referência uma AU-3, que atribui 783 octetos (87 colunas ×9 linhas) para a transmissão do VC-3.

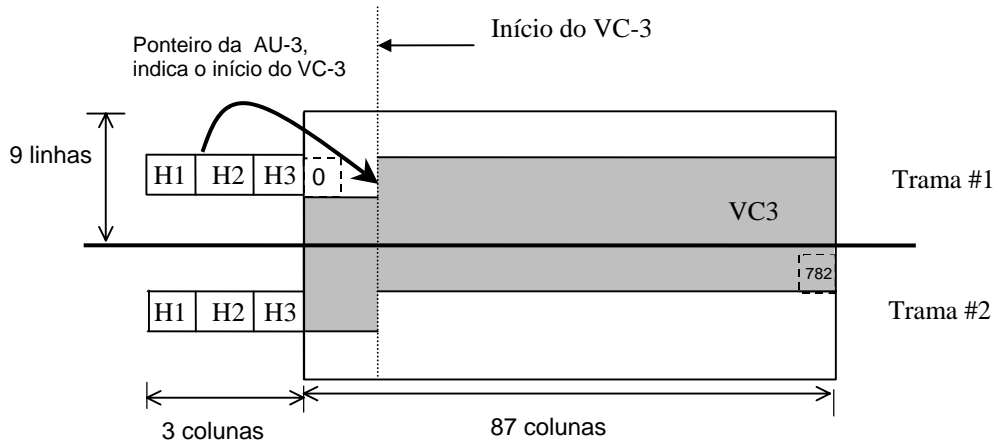



Figura 5.24 Exemplo da flutuação de um VC-3 num AU-3.

A informação do valor do ponteiro é transmitida nos octetos H1 e H2, enquanto o octeto H3 é usado para funções de justificação, como será visto posteriormente. O valor do ponteiro, correspondente ao deslocamento entre a posição 0 e a posição actual do octeto J1, é transportado como um número binário usando parte da palavra de 16 bits correspondente aos octetos H1 e H2. Como esse número poderá variar entre 0 e 782 são necessários 10 bits para o codificar. Os bits usados para esse fim são os dois bits menos significativos do octeto H1 e os oito bits do octeto H2, como se pode ver a partir da Figura 5.25. Os 4 bits mais significativos do octeto H1 (NNNN) constituem o indicador de novos dados ou NDF (*New Data Flag*). Quando o valor do ponteiro é introduzido inicialmente, ou quando é necessário fazer uma mudança não unitária do valor do ponteiro (ex: ressincronização da trama) tem-se NDF=1001. Neste caso, o novo valor do ponteiro deve ser recebido três vezes consecutivas, para o valor antigo ser substituído. Em condições de funcionamento normal tem-se NDF=0110. Os dois bits restantes são conhecidos por bits S e indicam o tipo de contentor virtual (Ex: SS=10, para o VC-3).

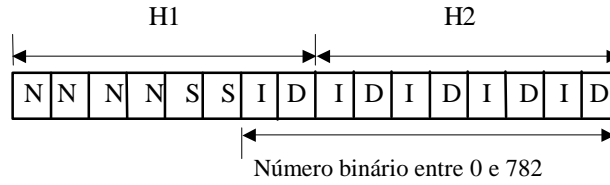



Figura 5.25 Formato do ponteiro.

Os 10 bits usados para transmitir a informação do ponteiro constituem duas palavras de cinco bits cada uma. Uma é a palavra I e a outra é a palavra D. As palavras I e D servem para incrementar ou decrementar o valor do ponteiro, e em conjunto com o octeto H3 são usadas para efeitos de justificação de trama.

Como já se referiu, devido ao contínuo processamento dos cabeçalhos podem surgir discrepâncias entre a fase da trama e a fase do contentor virtual. Pode-se ter assim, por exemplo, uma situação em que o contentor virtual flui temporariamente a um ritmo superior ao da trama (AU-3 no presente exemplo). Este problema é resolvido na SDH usando o octeto H3 para transportar um octeto do contentor virtual (**justificação negativa**) e diminuindo o valor do ponteiro de 1, correspondente a 1 octeto. A diminuição do valor do ponteiro é precedida da inversão dos bits D, enquanto o valor de NDF permanece igual a 0110. Em alternativa, quando temporariamente o contentor virtual se apresenta mais lento é enviado 1 octeto sem informação, ou seja em branco (**justificação positiva**) a seguir a H3 e o valor do ponteiro é aumentado de 1, correspondente a 1 octeto. Nesta situação, o incremento do ponteiro é antecedido da inversão dos bits I e da manutenção do NDF no seu estado normal.

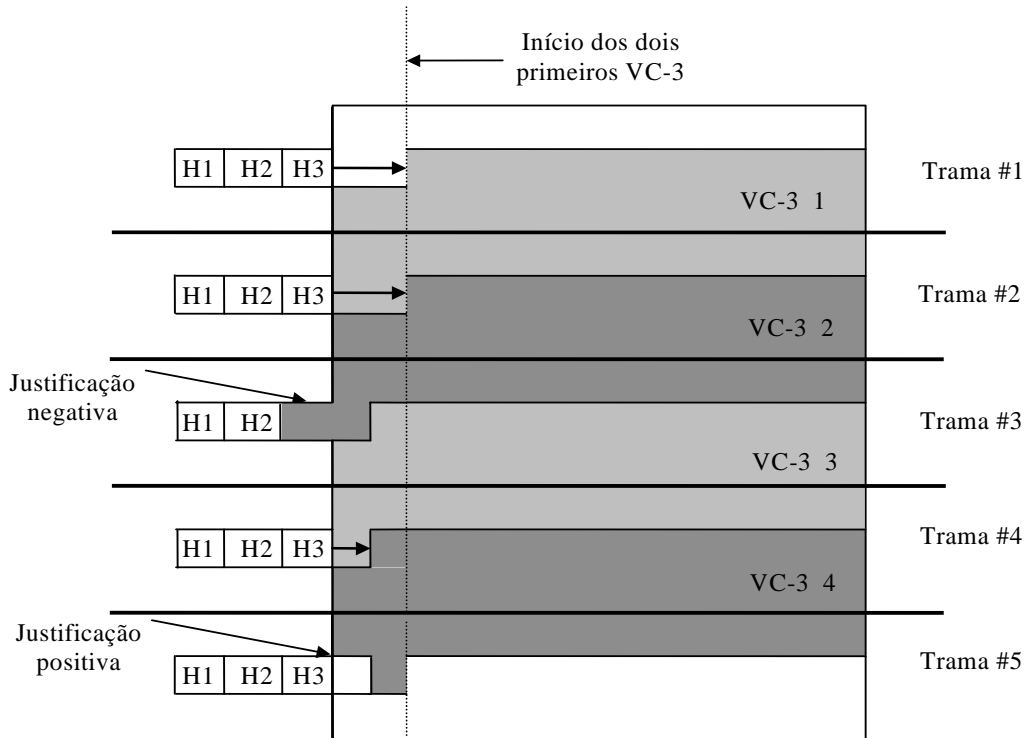



Figura 5.26 Mecanismo de justificação positiva e negativa para o caso da AU-3.

Na Figura 5.26 ilustra-se o mecanismo de justificação positiva e negativa considerando uma sequência de cinco AU-3 em cinco tramas seguidas.

A descrição efectuada sobre o papel do ponteiro AU-3 é, em termos gerais, válida para o ponteiro da AU-4. A diferença mais saliente reside no facto do ponteiro da AU-4, que é identificado, normalmente, como o ponteiro do sinal STM-1 (ver Figura 5.23), ter 9 octetos. Neste caso, têm-se três octetos H3 para efeitos de justificação, mas a estrutura da componente do ponteiro usada para indicar a posição do contentor virtual é exactamente igual à estrutura usada no ponteiro da AU-3, ou seja, é constituída pelos octetos H1 e H2. Os octetos X e Y são octetos com um valor invariável (X=11111111, Y=10011011). Como no caso da AU-4 se têm 2349 (261×9) octetos e o ponteiro só pode endereçar 783 posições, um incremento ou decremento do ponteiro de uma unidade, corresponde a uma deslocação do contentor virtual de 3 octetos.

5.3.8 Transporte dos contentores de ordem superior

Como a Figura 5.23 indica, uma trama STM-1 pode suportar uma AU-4 ou três AU-3. A AU-4 pode ser, por exemplo, usada para transportar um canal E4 (139.264 Mb/s), enquanto a AU-3 pode ser usada para transportar um DS3 (44.736 Mb/s), ou um E3 (34.368Mb/s), ou conjuntos de canais de débitos mais baixos.

Para uma melhor compreensão do processo de multiplexagem será útil analisar com algum detalhe a formação de algumas AUs. Vai-se começar por analisar a formação de uma AU-4 a partir de um sinal E4. A primeira fase consiste na formação do contentor virtual VC-4, a qual se designa por mapeamento. Como a Figura 5.28 mostra, o sinal E4 é inserido inicialmente num contentor C-4. Para adaptar os débitos do sinal E4 ao do contentor referido (149.76 Mb/s) é necessário inserir octetos de enchimento, ou seja, octetos sem informação. Após esta operação, é efectuada a adição do cabeçalho de caminho (POH) ao contentor formando-se, assim, o contentor virtual VC-4 de débito 150.336 Mb/s.

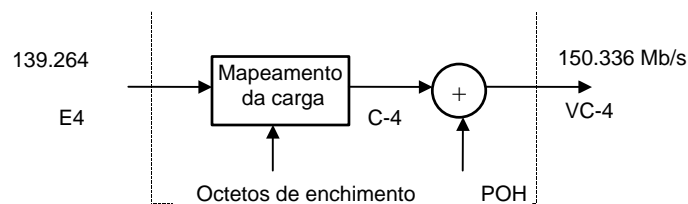



Figura 5.27 Processo de montagem do VC-4.

Após a obtenção do VC-4 é necessário enquadrá-lo na estrutura síncrona de transmissão através do uso de um ponteiro do tipo indicado anteriormente (ponteiro da AU-4) construindo a AU-4 (ver Figura 5.24), que é idêntica a um AUG. Neste ponto tem-se uma sequência de dados síncrona e pode-se adicionar o cabeçalho de secção para gerar a trama STM-1.

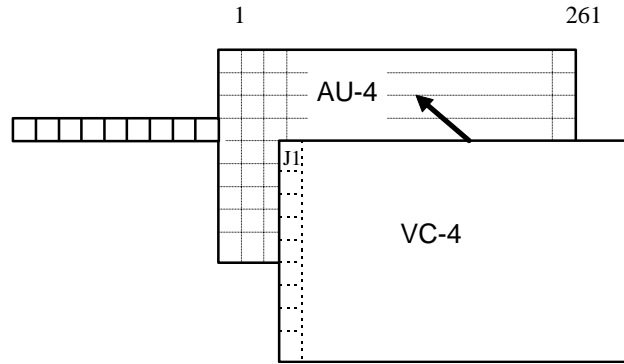



Figura 5.28 Estrutura do VC-4 e da sua inserção na AU-4.

O VC3 é o contentor virtual usado para transportar uma trama DS3, ou uma trama E3. A sua capacidade (carga+cabeçalho) é equivalente a 48.96 Mb/s. Este contentor pode ser representado por um agrupamento rectangular de octetos com 85 colunas e 9 linhas, cuja localização na AU-3 é indicada pelo ponteiro da AU-3. Daquelas colunas, uma é usada para o cabeçalho de caminho e as outras para a carga do cliente. A unidade administrativa AU-3 tem uma capacidade disponível para a carga de 87 colunas, já que como se conclui da Figura 5.24 a capacidade de transporte de um AUG (261x9) é equivalente à capacidade de três AU-3. Tem-se, assim, duas colunas a mais, que são preenchidas por octetos sem informação. A estrutura da AU-3 é apresentada na Figura 5.30 indicando-se, também, os três octetos do ponteiro da unidade administrativa. Três AU-3 são, em seguida, multiplexadas por interposição de octeto de modo a originar um AUG e após adicionar o cabeçalho de secção obtém-se o sinal STM-1.

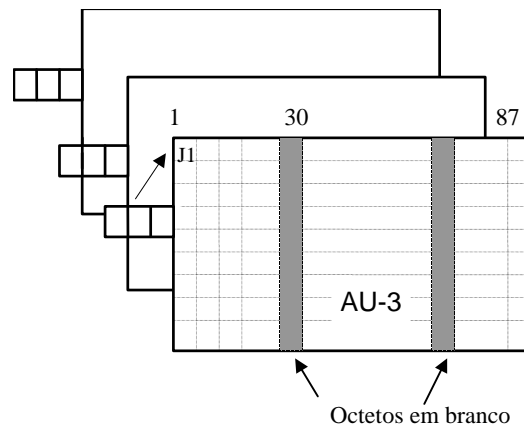



Figura 5.29 Estrutura do AU-3.

Note-se que, a SDH permite, também, transportar sinais com uma capacidade superior ao VC-4, através de um procedimento designado por **concatenação.** Segundo este procedimento, a carga digital correspondente a várias AU-4 consecutivas é amarrada, através da atribuição de um valor fixo (valor de concatenação) a todos os ponteiros das AU-4 do conjunto, com excepção do ponteiro da primeira. Esta estratégia permite sincronizar todas as AU-4 do conjunto, de modo que, qualquer ajuste no ponteiro da primeira AU-4 é replicado em todas as outras AU-4. De modo semelhante, o ponteiro do

cabeçalho de caminho só é transportado na primeira AU-4 e as colunas correspondentes a esse cabeçalho nas restantes AU-4 são preenchidas com octetos sem informação. Um conjunto de x AU-4 concatenadas designa-se por AU-4-xc. O valor de x igual a 4 é de particular interesse, já que a AU-4-4c, que é transportado numa trama STM-4, é usada para o transporte de tráfego ATM.

5.3.9 Transporte dos contentores de ordem inferior

No caso em que o VC-3 é considerado como um contentor virtual de ordem inferior tem-se uma outra alternativa para o seu transporte, fazendo uso da unidade tributária TU-3 e do contentor virtual VC-4. A estrutura bidimensional de uma TU-3 está ilustrada na Figura 5.30. Os três octetos do ponteiro (que aqui se designa por ponteiro de unidade tributária) H1, H2, e H3 têm, exactamente, as mesmas funções e modo de operação que o ponteiro da AU-3, embora aqui a disposição seja diferente. Também a gama dos valores do ponteiro é menor, ou seja, 765 (em vez de 783 da AU-3) e, tal como no caso da AU-3, o valor do ponteiro indica a posição do octeto da TU-3 onde o VC-3 começa (que corresponde ao octeto J1). O TUG-3 é obtido a partir da TU-3 preenchendo os últimos seis octetos da primeira coluna com octetos em branco.

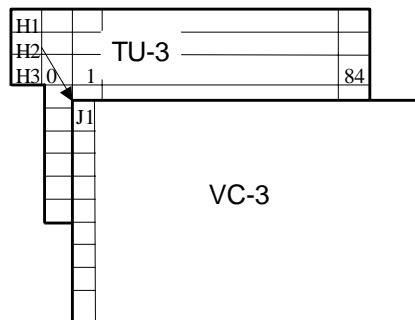



Figura 5.30 Estrutura da unidade tributário TU-3.

O TUG-3 pode ainda ser obtido por multiplexagem por interposição de octeto de sete TUG-2, e ao mesmo tempo preenchendo com octetos em branco as duas primeiras colunas do TUG-3, já que as restantes 84 colunas são suficientes para transportar toda a informação dos TUG-2.

Como o transporte dos contentores C-11, C-12 e C-2 na trama STM-1 (recorde-se que são contentores de ordem inferior) é significativamente diferente do transporte do C-3 e C-4, será conveniente, começar-se por analisar todo o processo associado ao transporte do C-12. Como se conclui a partir da Figura 5.20, o C-12 é usado para mapear um canal E1 a 2.048 Mbit/s. Como se sabe, um canal E1 corresponde à transmissão de 32 octetos todos os 125 μs, um valor ligeiramente inferior aos 34 octetos que correspondem ao C-12. Os dois octetos adicionais são usados como reserva e para funções de justificação similares às descritas aquando do estudo da formação das hierarquias plesiócronas de ordem superior. A operação seguinte consiste em gerar o cabeçalho de caminho de ordem inferior e adicioná-lo ao C-12, de modo a formar o VC-12. Este cabeçalho permanece ligado ao C-12 desde o ponto onde foi gerado, até ao ponto onde a carga de 2 Mbit/s

deixa a rede SDH, e permite identificar o C-12, monitorizar os erros, etc. O cabeçalho de caminho de ordem inferior é constituído por quatro octetos (V5, J2, N2 e K4). Devido à pequena dimensão do contentor C-12, não é eficiente concentrar todos os quatros octetos num único contentor virtual. Uma solução mais eficaz consiste em distribuí-los por quatro contentores virtuais, formando uma superstrutura em multitrama, com a duração de 500 μs, como se representa na Figura 5.31.

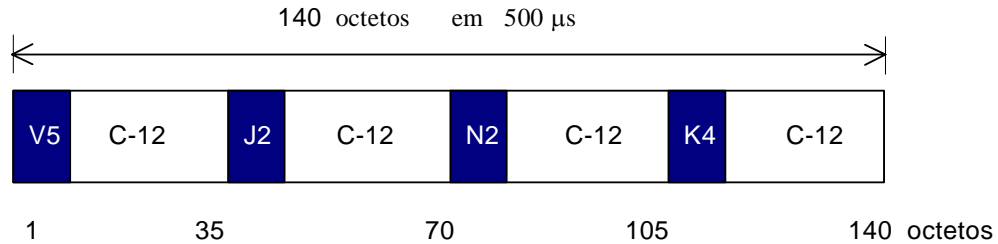



Figura 5.31 Estrutura de uma multitrama VC-12.

O octeto V5 sinaliza o início da multitrama VC-12. A Figura 5.32 ilustra a composição do octeto V5. Os primeiros 2 bits têm uma a função de monitorização de erros no próprio VC-12, através de um código BIP-2. O BIP-2 é calculado sobre todos os bits do conjunto de VCs que constituem a multitrama, incluindo os bits do cabeçalho de caminho e o resultado do cálculo é transmitido na multitrama seguinte. O bit seguinte (bit 3) é usado como indicador de erro distante (REI, *remote error indication*). Quando o código BIP-2 detecta um erro este bit é colocado a 1 e é enviado para o ponto de formação do VC-12 através de uma trama de retorno. O bit 4 é usado como um indicador remoto de falhas (RFI, *remote failure indication*) no caminho. Uma falha é definida como um defeito que se estende para além do tempo máximo permitido pelos mecanismos de protecção (cerca de 50 ms). No caso da detecção de uma falha este bit é colocado a 1. Os três bits seguintes (5,6,7) funcionam como etiqueta de sinal, indicando o tipo de mapeamento associado ao tributário em questão (não há transporte de carga, assíncrono, sincronizado em bit, sincronizado em octeto, etc). O último bit é um indicador de defeito distante (RDI, *remote defect indication*) e funciona como um alarme remoto, que é enviado para o elemento de rede onde o caminho é originado, quando por qualquer razão os dados não são entregues no destino.

Figura 5.32 Formato do octeto V5.

O octeto J2 é usado para validar o caminho de ordem inferior (traço de caminho de ordem inferior), funcionalidade idêntica à desempenhada por J1 para o caminho de ordem superior, ou seja, é usado como canal a 64 kb/s para enviar ciclicamente para a terminação do caminho uma determinada mensagem, de modo a verificar a integridade do caminho em questão. O octeto N2 é usado para funções de supervisão nas ligações em

cascata, enquanto o K4 é usado como um canal com funções de comutação de protecção automática (similar ao K1, K2, e K3).

O contentor virtual VC-12 é transportado numa unidade tributária TU-12. Para isso, é necessário adicionar o ponteiro de unidade tributária constituído pelos octetos V1, V2, e V3. Estes octetos têm funções semelhantes aos H1, H2 e H3 do ponteiro de unidade administrativa AU-3, ou seja, o V1 e V2 indicam o início do contentor virtual VC-12, apontando para o octeto V5, enquanto o V3 é usado para a operação de justificação negativa. A unidade tributária TU-12 também é estruturada numa multitrama com a duração de 500 μs. Essa multitrama está representada na Figura 5.33, onde o octeto V4 não tem funções normalizadas. Outra representação usual para essa multitrama consiste num conjunto de quatro quadros cada um dos quais com quatro colunas e nove linhas.

Multiplexando por interposição de octeto três TU-12 obtém-se um TUG-2 e agregando sete TUG-2 obtém-se um TUG-3, como já se referiu anteriormente. Como um VC-4 suporta três TUG-3 (ver Figura 5.23), pode-se concluir que um canal STM-1 tem capacidade para transportar 63 tributários E1.

Os contentores C-11 e C-2 têm, respectivamente, uma capacidade de 25 e 106 octetos e o seu processo de transporte é similar ao do C-12. Contudo, neste caso, um TUG-2 só permite alocar um TU-2 ou quatro TU-11.

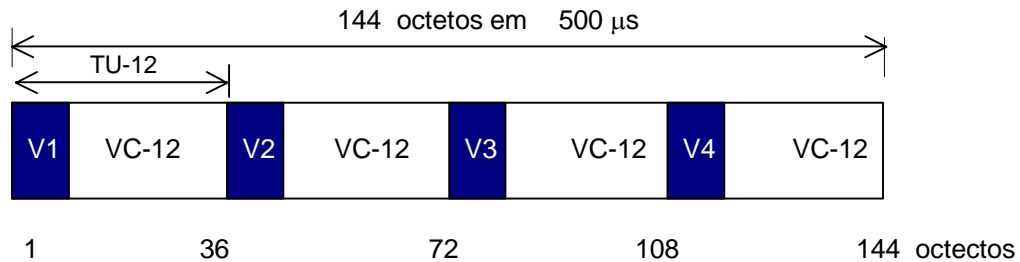



Figura 5.33 Estrutura de uma multitrama TU-12.

5.3.10 Aspectos de protecção

Como já se referiu as funções de suporte operacional dos sistemas plesiocrónos são muito limitadas, enquanto, em alternativa, os sistemas SDH dispõem de um cabeçalho com capacidade suficiente para monitorizar o estado da rede, trocar informação entre os elementos da rede e realizar funções de recuperação da rede, em caso de falhas, sem necessidade de intervenção humana.

Por sua vez, devido às elevadas capacidades das redes SDH, qualquer falha na rede (corte nas fibras ou defeito nos nós) pode ter consequências desastrosas para os clientes da rede de transporte. Assim, nas redes SDH são exigidos tempos de recuperação de falhas da ordem das dezenas de milissegundos (tipicamente 50 ms), em contraposição aos tempos de recuperação de várias horas típicos das redes PDH. Estes objectivos podem ser atingidos usando uma filosofia de auto-recuperação sem intervenção humana, fazendo

uso das potencialidades da tecnologia SDH. [65]Diz-se que uma rede SDH tem capacidade para auto-recuperação quando em caso de falha num caminho, ou numa secção, o tráfego é comutado automaticamente para um caminho ou secção alternativa.

As redes SDH (quer de acesso, quer dorsais) usam normalmente uma topologia física em anel como uma solução atractiva para garantir auto-recuperação. Esses anéis usam ADMs em cada um dos nós e podem ser bidireccionais ou unidireccionais como ser irá ver nas secções seguintes.

5.3.10.1 Anel unidireccional com protecção a nível de caminho

Na Figura 5.34 representa-se um anel unidireccional com protecção realizada através de comutação de caminho. Nesse anel, os diferentes nós estão ligados por duas fibras ópticas, uma para função de protecção e outra de serviço. Cada nó proporciona ainda a função de inserção/extracção, ou seja é constituído por um ADM.

O anel diz-se unidireccional porque o tráfego flui na fibra de serviço num único sentido. Atendendo à Figura 5.34 (a) o tráfego de serviço de A para C flui no sentido dos ponteiros do relógio, enquanto o tráfego de C para A também flui no mesmo sentido. Note-se, no entanto, que os sinais transmitidos são replicados e vão alimentar a fibra de protecção, onde são transmitidos em sentido contrário (protecção 1+1). No estado de funcionamento normal, o sinal extraído no nó destinatário corresponde ao sinal que é recebido através da fibra de serviço.

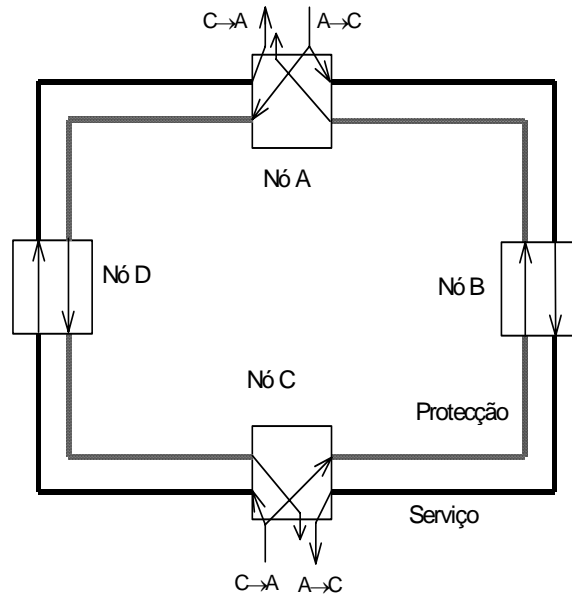



Figura 5.34 (a) Anel unidireccional com protecção a nível de caminho em estado normal.

[65] Com as facilidade oferecidas pelos esquemas de protecção da SDH é possível construir redes de transporte com uma fiabilidade extremamente elevada, sendo possível garantir, normalmente, a chamada fiabilidade dos cinco "9s", ou seja 99,999 % de disponibilidade (5 minutos de falha num ano).

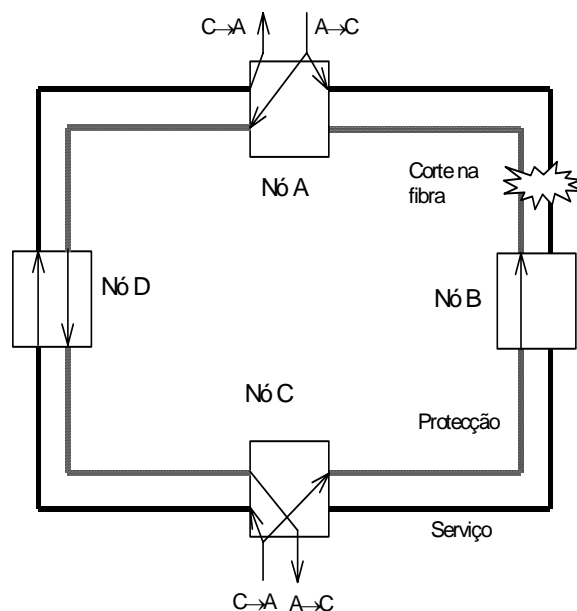



Figura 5.34 (b) Anel unidireccional com protecção a nível de caminho em estado de protecção.

A qualidade do sinal recebido correspondente a um determinado caminho é monitorizada permanentemente através do cabeçalho de caminho. Quando o sinal se degrada, e a qualidade desce abaixo do mínimo exigido, o serviço é comutado pelo nó destinatário para o sinal proveniente da fibra de protecção. Este tipo de comportamento é exemplificado na Figura 5.34 (b), onde um corte na fibra, no arco (segmento) do anel entre A e B, leva a que o nó C passe a extrair o tráfego proveniente de A a partir da fibra de protecção. Como este tipo de protecção não envolve nem comunicação entre nós, nem conhecimento da configuração do anel, torna possível que a recuperação do tráfego seja muito rápida, normalmente inferior a 50 ms.

A principal desvantagem desta técnica de protecção prende-se com o facto de se exigir ao anel capacidade para escoar todo o tráfego agregado pelos diferentes nós, já que o tráfego de qualquer tributário que entra na rede através de qualquer nó terá necessariamente de percorrer todo o anel. Por exemplo, no caso exemplificado na Figura 5.34, se cada ADM tem capacidade para inserir/extrair um STM-1, o anel deverá ter uma capacidade STM-4, o que implica a utilização de ADMs STM-4, de modo a vazar todo o tráfego do anel.

5.3.10.2 Anel unidireccional com protecção a nível de secção

Na Figura 5.35 representa-se um anel unidireccional com protecção a nível de secção. Os nós estão interligados de modo idêntico ao caso anterior, mas no estado de operação normal só a fibra de serviço transporta tráfego, enquanto a fibra de protecção funciona como reserva (protecção 1:1). Como se pode observar, o caminho estabelecido entre os nós A e C atravessa todo o anel no sentido dos ponteiros do relógio, de modo que, este sistema tem as mesmas limitações de capacidade que o anterior.

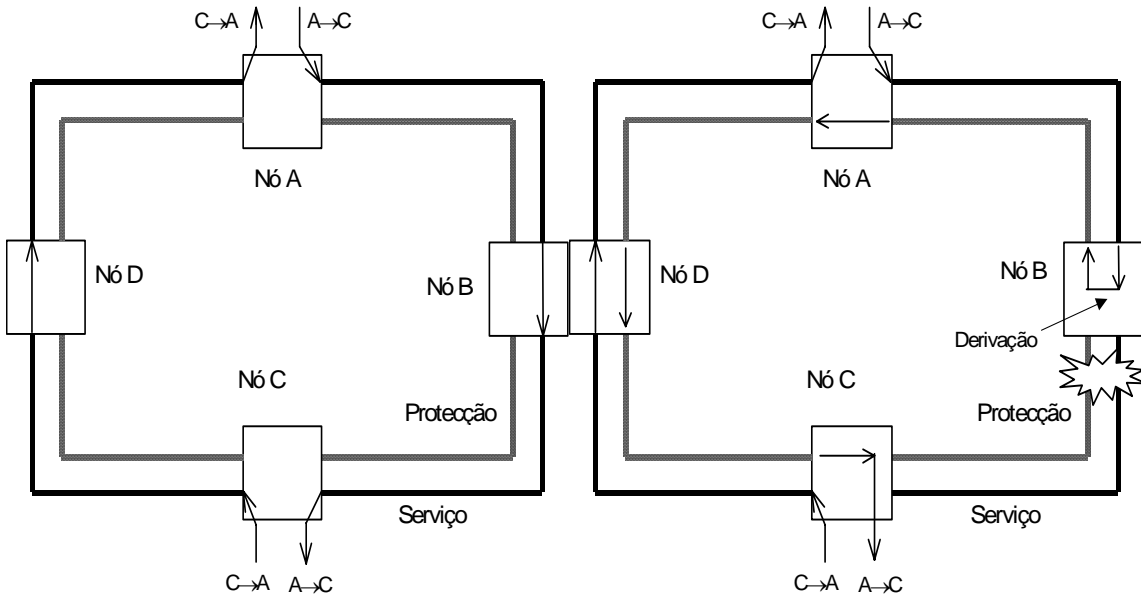



Figura 5.35 Anel unidireccional com protecção a nível de secção: a) estado normal; b) estado comutado.

Neste tipo de anel a análise do desempenho da transmissão é feito a nível do cabeçalho de secção. No caso do corte de um cabo, o tráfego é interrompido, sendo a falha indicada a nível do cabeçalho de secção. Depois de detectada a falha inicia-se o processo de restauro fazendo uso da comutação de protecção automática (APS: *automatic protection switching)*. O APS faz uso dos octetos K1 e K2 do cabeçalho de secção de multiplexagem, que são usados para transportar a informação de gestão de protecção. A recomendação G.841 do ITU-T indica que o protocolo APS para este tipo de anel ainda está em fase de estudo.

5.3.10.3 Anel bidireccional com duas fibras com protecção a nível de secção

Num anel bidireccional com duas fibras, os canais de serviço e de protecção usam a mesma fibra, reservando-se normalmente metade da capacidade total para protecção. Assim, num anel com capacidade STM-N, os módulos STM-N transferidos nos dois sentidos reservam as AU-4 numeradas de 1 a N/2 para o transporte do tráfego de serviço e os AU-4 numerados de N/2+1 a N para o transporte do tráfego de protecção.

Na Figura 5.36(a) representa-se um anel bidireccional em estado normal, onde metade da capacidade dos módulos STM-N que se propagam no sentido dos ponteiros do relógio (A→C) e em sentido contrário (C→A) é reservada para funções de protecção. No caso de um corte das fibras entre os nós B e C todo o tráfego é interrompido. A falha é indicada ao nível do cabeçalho de secção de multiplexagem e o processo de restauro é iniciado usando um mecanismo APS, que leva à comutação automática do tráfego para a capacidade reservada nos módulos STM-N (ver Figura 5.36b). Como esta comutação é feita a nível de *time-slots,* os ADMs usados nesta operação têm de apresentar uma funcionalidade TSI (*time-slot interchange*), a qual irá ser estudada no Capítulo 6.

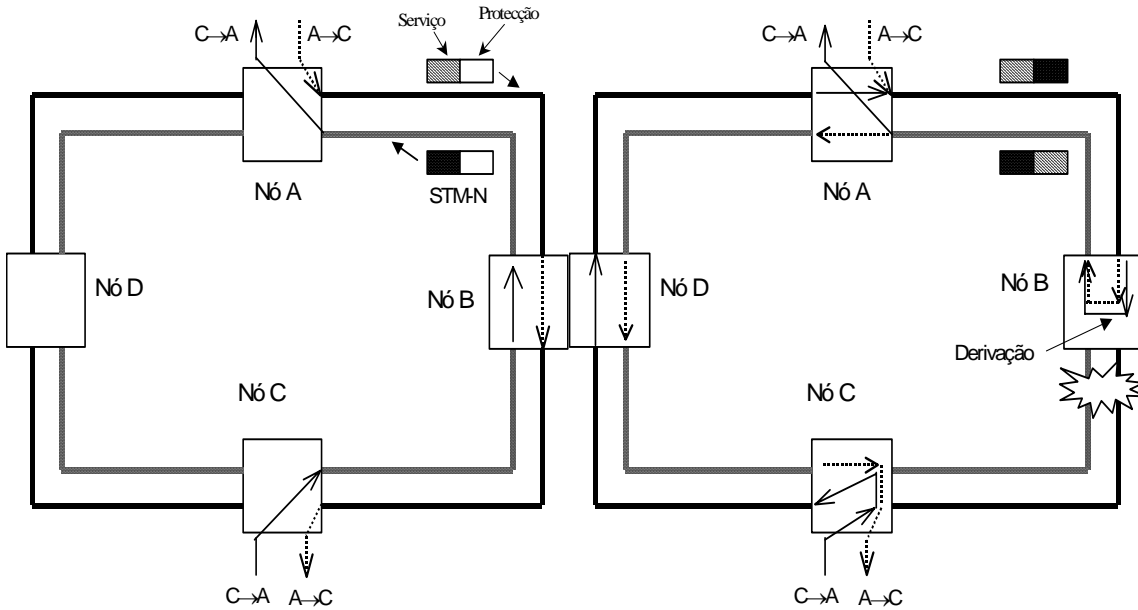



Figura 5.36 Anel bidireccional com duas fibras com protecção a nível de secção. (a) estado normal; (b) estado de protecção.

O protocolo APS aplicado no contexto dos anéis bidireccionais faz também uso dos octetos K1 e K2 do cabeçalho de secção. Os primeiros quatro bits (1 a 4) de K1 são usados para os sinais de comando associados à protecção (ex: estabelecimento de derivações), enquanto os quatro últimos bits (5 a 8) são usados para transmitir o número de identificação do nó destinatário (os nós do anel são identificados por um número escolhido entre 0 e 15, o que permite ter no máximo anéis com 16 nós). No que se refere ao octeto K2, os primeiros quatro bits (1 a 4) transportam o número de identificação do nó fonte e os quatro restantes informação adicional, como por exemplo o estado do nó onde o K2 foi originado (nó fonte).

O **estado de funcionamento normal** é caracterizado pelo facto de cada nó emitir em ambos os sentidos octetos K, não contendo nenhum pedido de intervenção (bits 1 a 4 de K1=0000). Quando qualquer nó detecta uma degradação do sinal recebido susceptível de desencadear os mecanismos de protecção, envia uma mensagem pedindo o estabelecimento de uma derivação nos dois sentidos de transmissão (percurso mais longo e mais curto). A mensagem inclui o número de identificação do nó fonte e do nó destinatário. O nó destinatário é o nó adjacente ao nó fonte do lado oposto à via com problemas. Os nós intermediários deixam passar a mensagem sem alterar o conteúdo dos octetos K. Quando o pedido de estabelecimento de uma derivação é recebido pelo nó destinatário, este passa para o **estado comutado** (derivação estabelecida) e reencaminha o tráfego afectado pela falha, pela capacidade reservada à protecção. A Figura 5. 36(b) mostra um anel no estado comutado em resposta a um corte nas fibras entre B e C.

Quando a falha é reparada, o sistema passa ao estado de operação normal, a derivação é removida e todo o tráfego passa a ser encaminhado pela capacidade de serviço.

5.3.10.4 Anel bidireccional com quatro fibras com protecção a nível de secção

Num anel bidireccional com quatro fibras, duas fibras são usadas para escoar os canais de serviço, enquanto as outras duas transportam os canais de protecção, como se mostra na Figura 5.37. No caso de corte nas fibras entre os nós B e C, o tráfego entre esses dois nós é interrompido. A falha é indicada ao nível do cabeçalho de secção de multiplexagem e o processo de restauro inicia-se usando um processo APS do tipo descrito anteriormente. Através deste processo, estabelecem-se derivações entre as fibras de serviço e as fibras de protecção nos nós adjacentes ao ponto onde ocorreu o corte, o que permite restaurar o tráfego. O anel com quatro fibras tem ainda possibilidade de restaurar o tráfego através de um outro processo, designado por comutação de arco, nos casos em que os cortes só afectam uma ou as duas fibras de serviço, ou em caso de falhas dos emissores e receptores associados a essas fibras. Qualquer que seja o tipo de falha esta é indicada através do cabeçalho de secção, e o tráfego no percurso afectado é comutado para um percurso protegido nas extremidades do arco (segmento do anel). Esta comutação não afecta os outros arcos, e a comutação num arco pode ocorrer independentemente da comutação nos outros.

Uma desvantagem desta técnica de protecção está associada ao facto de requerer dois ADMs em cada nó. Estes dois ADMs podem ser integrados num único ADM (designado por ADM 1+1), que termina as quatros fibras e dispõe de um sistema de controlo único. A opção por qualquer uma destas soluções resulta de um compromisso entre custo e fiabilidade. A solução baseada no anel com dois ADMs independentes é mais cara, mas oferece um maior grau de fiabilidade relativamente a falhas nos ADMs, sendo por isso a solução que oferece o maior grau de sobrevivência. A solução ADM 1+1 é mais barata, mas não oferece protecção para falhas a nível do ADM.

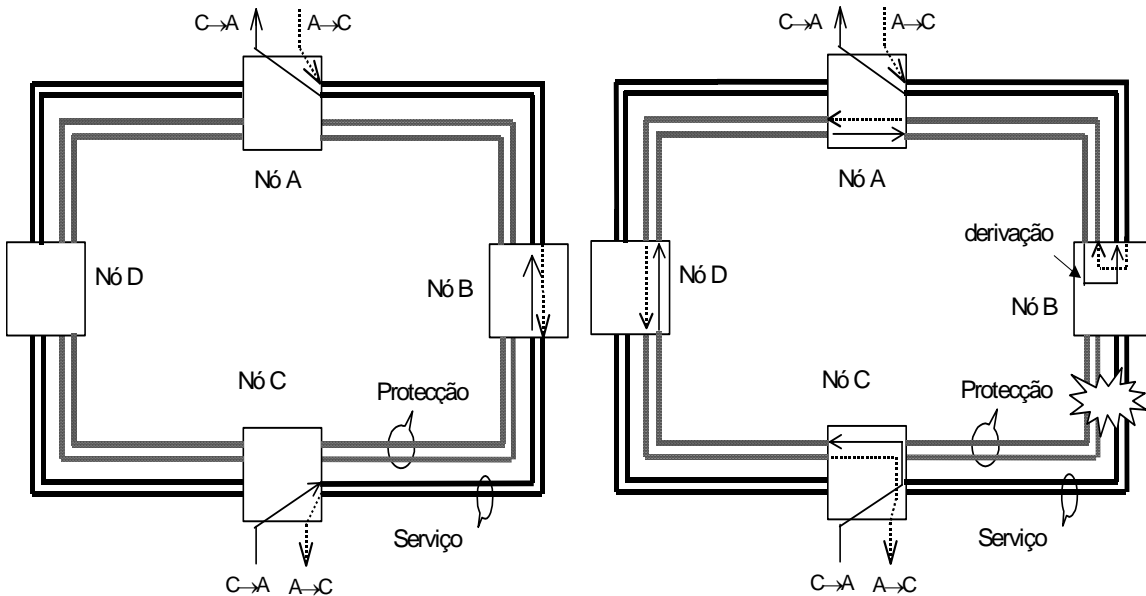



Figura 5.37 Anel bidireccional com quatro fibras com protecção a nível de secção.
(a) Estado normal; (b) Estado comutado.

Será ainda de referir que o anel descrito poderá ainda ser implementado usando só duas fibras e recorrendo a técnicas WDM para simular um anel com quatro caminhos ópticos, ou seja transmitindo dois comprimentos de onda por fibra.

5.4 Problemas

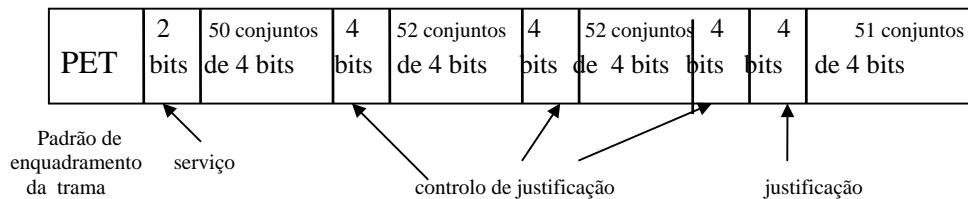
5.1 a) Admita que um multiplexer TDM tem na sua entrada 4 canais, cada um com um débito nominal de 2.048 Mb/s, e que a flutuação máxima esperada nesse débito é de 10 kb/s. Admitindo que cada trama pode incluir apenas um bit de justificação para cada um dos quatro canais, determine o comprimento máximo da trama (excluindo os bits de controlo de justificação+sincronismo). Assuma, que a duração da trama de entrada é idêntica à da trama de saída.

b) Considere que a justificação de bit para cada um dos canais é indicada em cada trama pelo conjunto de bits C_{in}, i=1,...4, n=1,....,N. Pretende-se que o tempo médio entre erros (em relação aos bits de justificação) seja superior a 100 horas. Determine qual o valor limite da probabilidade de errar um bit para que seja atingido esse limite: (i) apenas com um bit para indicar justificação i.e., N=1; (ii) usando três bits para indicar justificação, i.e, N=3. Faça as hipóteses que considerar necessárias para resolver o problema.

5.2 Considere um sistema TDM da primeira hierarquia.

a) Desenhe um circuito que permita reconhecer o padrão de enquadramento da trama.

b) Considere os seguintes padrões de enquadramento de trama 0011011 e 1110010. Admitindo que pretendia minimizar o grau de semelhança entre o padrão e as suas versões deslocadas, qual seria o padrão que escolhia.

5.3 A trama da 2ª hierarquia plesiócrona do CCITT tem a seguinte estrutura:

O padrão de enquadramento de trama é o seguinte: 1111010000. Considere que há perda de enquadramento da trama, quando são recebidos 4 padrões de enquadramento errados. Por sua vez o enquadramento é considerado recuperado, quando se detectam 3 padrões de enquadramento correctos. Tendo presente que a probabilidade de erro de bit é 10^{-4}, determine:

a) O número de bits/trama;

b) O número de bits/tributário;

c) Taxa máxima de justificação/tributário;

d) Probabilidade de declarar indevidamente a perda de enquadramento da trama;

e) Tempo médio entre declarações indevidas de perda de enquadramento de trama;
f) Probabilidade de decisão errada sobre a justificação;
g) Tempo médio de aquisição do enquadramento de trama.

5.4 Uma célula ATM ocupa 53 octetos. Quantas células ATM podem ser transportadas num VC-4.

5.5 Descreva o processo associado ao empacotamento de um sinal E4 (140 Mb/s) numa trama STM1 da hierarquia digital síncrona. Determine o número de sinais E1 (2 Mb/s) que são transportados num sinal CEPT4. Compare esse número com o número de sinais E1 que é possível tansportar na trama STM1 usando contentores C12 e tire conclusões.

5.6 Determine o valor máximo e o valor mínimo do débito binário de um sinal VC-3, que pode ser transportado numa AU-3. Determine ainda o número de octetos em branco inseridos no processo de formação do contentor C-3. Admita que o C-3 é usado para transportar um DS3 (44.736 Mbit/s).

5.7 Determine o valor máximo e o valor mínimo do débito binário de um sinal VC-4, capaz de ser acomodado numa AU-4. Determine também o número de octetos em branco inseridos no processo de formação do contentor C-4, o qual é usado para transportar um E4.

5.8 Represente num diagrama 4 tramas consecutivas STM-1 nas quais tem lugar uma acção de justificação negativa a nível da AU-4, considerando que na trama inicial o valor do ponteiro é igual a 87. Especifique o valor decimal do ponteiro (10 bits I, D) e marque a posição do VC-4 e do seu cabeçalho de caminho em todas as tramas.

5.9 Admita que um sinal STM-1, que transporta um contentor virtual VC-4 é transmitido num meio que induz uma taxa de erros de bit de 10^{-9}. Pretende-se que:
a) Indique o código BIP usado para monitorizar o sinal STM-1 a nível de secção de multiplexagem e do contentor VC-4.
b) Determine o desempenho do VC-4 expresso em ES (segundos errados) e BE (blocos errados) considerando que o intervalo de medida é de 15 minutos.

5.10 Considere um anel SDH unidireccional de duas fibras com 5 nós (ADMs) que usa protecção a nível de secção. Admita que a topologia lógica desse anel é uma topologia em malha, onde em cada nó é inserida/extraído para cada um dos outros nós um E3.
a) Explique como esse anel se comporta em presença de uma falha na fibra.
b) Calcule a capacidade dos ADMs usados na rede.
c) Responda às alíneas anteriores considerando em alternativa um anel bidireccional com 4 fibras com protecção a nível de secção.

5.5 Referências

[Fl97] J. E. Flood, *Telecommunication Networks*, The Institution of Electrical Engineers, 1997.

[Sc81] M. Schwartz, *Information, Transmission, Modulation, and Noise, McGraw-Hill*, 1981.

[Se97] M. Sexton, A. Reid, *Broadband Networking: ATM, SDH, and SONET,* Artech House, 1997.

[Pe95] J. Perdigoto, *Análise de Desempenho em Redes de Transporte Baseadas na Hierarquia Digital Síncrona,* Tese de Mestrado, Universidade de Coimbra, 1995.

[Le95] W. S. Lee, D. C. Brown, *Advances in Telecommunications Networks,* Artech House, 1995.

[Wu92] Tsong-Ho Wu, *Fiber Network Service Survivability*, Artech House, 1992.

[Bo97] G. Bouyer, *Les réseaux synchrones étendus PDH et SDH*, Hermes, Paris 1997.

6 Tráfego, comutação e sinalização

6.1 Fundamentos da teoria do tráfego

6.1.1 Introdução

O objectivo de uma rede telefónica de comutação de circuitos consiste em estabelecer **circuitos** (ligações bidireccionais) extremo a extremo envolvendo a afectação de vários recursos da rede (vias de transmissão, sistemas de multiplexagem, centrais de comutação, etc.) durante todo o período de duração de uma chamada. Assim, do ponto de vista da rede não é o fluxo de informação nos circuitos individuais que é importante, mas sim a afectação dos referidos recursos. Ora, como o número de chamadas telefónicas oferecidas a uma rede, bem como a duração dessas chamadas variam aleatoriamente, a análise e dimensionamento de uma rede telefónica terá necessariamente de envolver o recurso a modelos estatísticos apropriados. A **engenharia de teletráfego** é dedicada exactamente ao estudo desses modelos.

Uma das aplicações mais relevantes da engenharia de teletráfego consiste em dimensionar o número de troncas (circuitos) necessárias numa determinada ligação entre centrais telefónicas. Por razões económicas não é concebível projectar uma rede telefónica com equipamento suficiente para dar resposta imediata a todos os pedidos de serviço (chamadas), sendo aceitável um certo bloqueio (pedidos de serviço recusados) ou atraso (nas redes de dados). Por conseguinte, aspectos como modelação e previsão de tráfego telefónico serão de especial importância no dimensionamento referido.

6.1.2 Características do tráfego telefónico

Se o número de chamadas em progresso num sistema de telecomunicações de grandes dimensões, tal como uma central telefónica, ou uma via de transmissão, fosse registado durante um intervalo de tempo de poucos minutos seria de esperar um andamento do tipo representado na Figura 6.1. A Figura 6.2 representa uma variação (hora-a-hora) das chamadas durante um dia de trabalho típico, numa central telefónica servindo uma zona mista (residencial +empresarial). Existem poucas chamadas durante a noite. O número de chamadas cresce à medida que as pessoas vão chegando ao trabalho atingindo um máximo a meio da manhã. Esse número cai por volta da hora do almoço e volta a aumentar da parte da tarde. Decresce novamente à medida que as pessoas regressam a casa, notando-se ainda um ligeiro pico ao princípio da noite em resultado de possíveis contactos sociais.

Pode-se ver que a hora em que ocorre um maior número de chamadas se situa entre as 10 e as 11 horas da manhã. Essa hora designa-se por **hora de ponta,** ou **hora mais carregada,** e é usada geralmente para analisar a actividade das redes telefónicas públicas. A utilização de medidas de tráfego na hora mais carregada para analisar e projectar as redes telefónicas resulta de um compromisso entre um projecto virado para a inteira

utilização (o que inclui as horas nocturnas com quase utilização nula) e um projecto virado para os picos de tráfego de curta duração, resultantes de fenómenos imprevisíveis (desastres naturais, etc.), e de eventos especiais e concursos na TV.

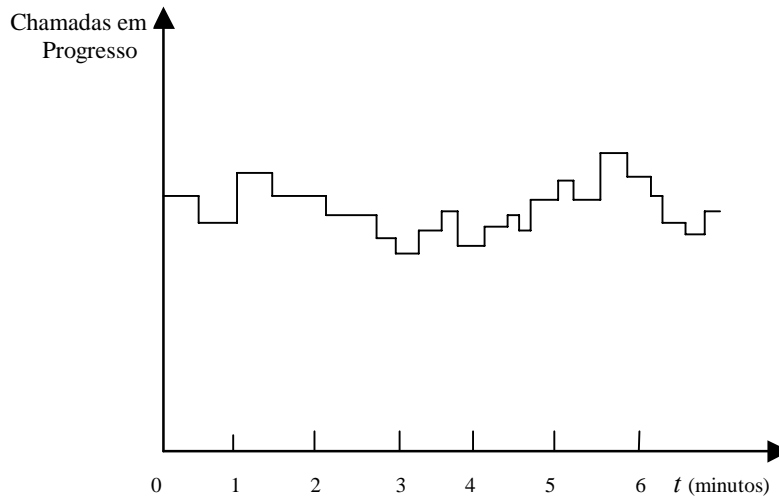



Figura 6.1 Variação do tráfego a curto prazo.

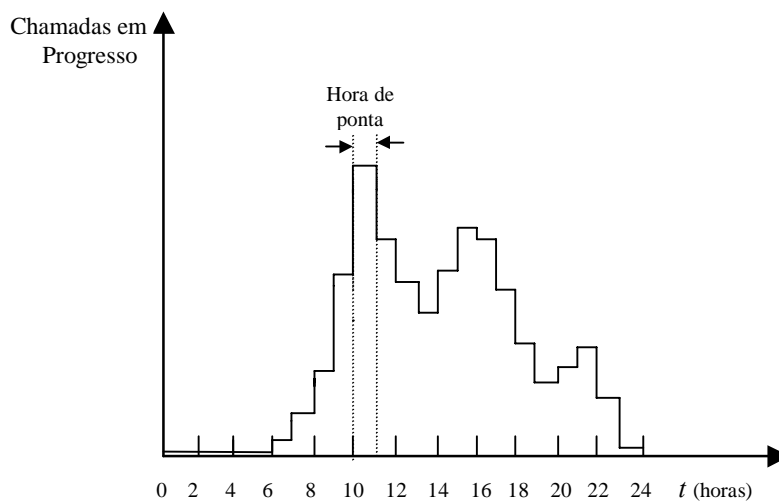



Figura 6.2 Variação do tráfego telefónico ao longo de um dia.

6.1.3 Medidas de tráfego

As centrais de comutação telefónica desempenham um papel fundamental no estabelecimento de um circuito entre dois quaisquer assinantes. No estudo do tráfego que se vai realizar admite-se que as centrais telefónicas são de acesso total, permitindo, assim, ligar qualquer entrada livre a qualquer saída livre. Na Figura 6.3 representa-se um modelo

genérico de uma central telefónica de acesso total com *M* linhas de entrada (troncas) e *N* linhas de saída (troncas). O estabelecimento de um circuito em resposta a um pedido da chamada implica apresar uma das *N* troncas de saída e ligá-la a uma das *M* linhas de entrada. Se todas as troncas de saída estão ocupadas a chamada é bloqueada. Obviamente para haver bloqueio é necessário que $M>N$, ou seja, o número de entradas é superior ao número de saídas. Um sistema com esta característica usa concentração e como permite uma partilha dos recursos de saída (troncas de saída) conduz a uma redução dos custos.

Figura 6.3 Modelo genérico de uma central de comutação de circuitos.

Seja $K_i(t)$ o **tráfego instantâneo** transportado pela tronca *i*, tendo-se $K_i(t)=1$ se a tronca está activa e $K_i(t)=0$, em caso contrário. Então o tráfego instantâneo transportado por um grupo de *M* troncas é dado por:

$$m(t)=\sum_{i=1}^{M}K_i(t). \tag{6.1}$$

A equação anterior define o número de troncas activas num determinado instante. O **volume de tráfego** escoado pelo grupo de troncas considerado durante o tempo *T*, não é mais do que a soma dos tempos de ocupação desse grupo no intervalo de tempo *T*, isto é

$$A_v=\int_0^T m(t)dt \tag{6.2}$$

A equação anterior pode-se simplificar se se admitir que o número médio de chamadas no intervalo de tempo *T* é igual a *C* e que a duração média dessas chamadas é igual a *h,* vindo

$$A_v=Ch \tag{6.3}$$

Uma medida de tráfego mais útil é a **intensidade de tráfego** (também designada por fluxo de tráfego). A intensidade de tráfego *A* é obtida dividindo o volume de tráfego pelo intervalo de tempo de medida, ou seja

$$A=\frac{A_v}{T}=\frac{Ch}{T}=\lambda h \tag{6.4}$$

onde λ representa o número médio de chamadas por unidade de tempo. Embora a intensidade de tráfego seja adimensional é normalmente expressa em Erlang (abreviatura E) em homenagem ao matemático dinamarquês A. K. Erlang, que estabeleceu as bases da teoria do tráfego. O tráfego máximo transportado por uma tronca será de 1 E, o que corresponde à tronca estar permanentemente ocupada durante o período de medida, normalmente uma hora. A taxa de chamadas na hora mais carregada é o número médio de chamadas realizadas por assinante durante a hora de ponta. Esta taxa varia entre 0.5 para uma região rural até perto de 1.5 para uma zona muito activa de uma grande cidade. Como a duração de cada chamada varia entre os 3 e 4 minutos, tem-se um tráfego por assinante compreendido entre 0.025 e 0.1 E.

Exemplo 6.1
Em média uma companhia faz 200 chamadas com duração média de 3 minutos e recebe 300 chamadas com duração média de 4 minutos. Pretende-se determinar 1) a intensidade do tráfego gerado ;2) a intensidade do tráfego recebido ;3) a intensidade do tráfego total.

Solução:
1. Intensidade do tráfego gerado: 200x3/60= 10 E
2. Intensidade do tráfego recebido: 300x4/60=20 E
3. Intensidade do tráfego total: 10 +20 =30 E

6.1.4 Congestionamento

Como já se referiu é antieconómico projectar uma rede de telecomunicações com equipamento suficiente para escoar simultaneamente todo o tráfego oferecido à rede. Assim, pode acontecer que todas as troncas de um determinado grupo de troncas estejam ocupadas e não possam aceitar mais chamadas. Este estado designa-se por **congestionamento**. Nas redes de comutação de pacotes todas as chamadas (mensagens) que chegam durante um período de congestionamento esperam numa fila até que uma tronca de saída fique livre. Por isso, essas chamadas são atrasadas mas não perdidas. Os sistemas com estas características designam-se por sistemas com **filas de espera.** Nos sistemas de comutação de circuitos, como é o caso das centrais telefónicas, todas as tentativas para realizar chamadas num sistema congestionado são infrutíferas. Estes sistemas designam-se por **sistemas com perdas.**

Num sistema com perdas o **tráfego transportado** é inferior ao **tráfego oferecido,** sendo a diferença o **tráfego perdido**. Define-se o **grau de serviço** B, como a relação entre o tráfego perdido e o tráfego oferecido na hora de ponta, ou seja

$$B = \frac{\text{Tráfego perdido}}{\text{Tráfego oferecido}} \tag{6.5}$$

Assim, se é oferecido um tráfego de A Erlangs a um grupo de troncas com um grau de serviço de B o tráfego perdido é AB e o tráfego transportado é A(1-B) Erlangs.
A escolha do grau de serviço resulta de um compromisso entre dois factores contraditórios. Se é muito elevado os utilizadores fazem muitas tentativas infrutíferas o

que é motivo de insatisfação. Se por outro lado é muito baixo tem-se despesas desnecessárias em equipamento que é raramente usado. Um objectivo razoável para as redes de telecomunicações modernas consiste em considerar um grau de serviço na hora de ponta que não ultrapasse 1%.

Definido o grau de serviço e o tráfego oferecido é necessário calcular o número de troncas requeridas. É este problema que irá ser analisado nas secções seguintes.

Exemplo 6.2
Durante a hora de ponta são oferecidas, a um grupo de troncas, 1000 chamadas telefónicas. Admitindo que são perdidas 10 e que a duração de cada chamada é de 3 minutos, determine: 1) grau de serviço ; 2) a duração total dos períodos de congestionamento.

Solução: Tráfego oferecido→A=1000x3/60=50 E
Tráfego perdido→ AB=10x3/60=0.5 E
Grau de serviço → B=0.5/50=0.01
Duração total dos períodos de congestionamento→0.01x3600=36 s

6.1.5 Modelação de tráfego

Com o objectivo de simplificar a análise admite-se que o tráfego oferecido a um determinado sistema de telecomunicações (por exemplo uma central telefónica) é puramente aleatório. Isto quer dizer que, os eventos associados à chegada e terminação das chamadas são eventos aleatórios e independentes. Um tráfego com estas características é conhecido por tráfego sem memória e implica que o número de fontes geradoras de chamadas seja muito elevado. Nesta situação, o número de chegadas de chamadas num determinado intervalo de tempo T é descrito por um **processo de Poisson**. Assim, a probabilidade $P(k)$ de chegarem k chamadas no intervalo de tempo T é descrita por

$$P(k) = \frac{e^{-\lambda T}(\lambda T)^k}{k!} \qquad k=0, 1, 2, 3\ldots \tag{6.6}$$

onde λ é o número médio de chamadas ocorridas por unidade de tempo. Atendendo às propriedades da distribuição de Poisson tem-se que o número médio de chamadas chegadas no intervalo de tempo T é dado por

$$\mathrm{E}(k)=\Lambda=\lambda T \tag{6.7}$$

enquanto a variância desse número é dada por

$$\sigma_k^2 = \lambda T \tag{6.8}$$

Os processos de Poisson apresentam duas propriedades fundamentais:

- **estacionariedade:** a probabilidade de um dado número de eventos (chegada de chamadas) ocorrerem num intervalo de tempo determinado depende apenas da duração desse intervalo e não do seu posicionamento no tempo;

- **incorrelação:** traduz independência entre eventos (chegada de chamadas) que ocorrem em intervalos de tempo disjuntos.

Exemplo 6.3
Em média chega uma chamada em cada 5 segundos. Durante 20 segundos qual é a probabilidade que:
1. Não chegue nenhuma chamada?
2. Chegue uma chamada?
3. Cheguem duas chamadas?
4. Cheguem mais de duas chamadas?

Solução:
$P(k) = \frac{e^{-\Lambda}\Lambda^k}{k!}$, onde $\Lambda=4$
1. $P(0)=0.018$
2. $P(1)=0.073$
3. $P(2)=0.147$
4. $P(>2)=1-P(0)-P(1)-P(2)=0.762$

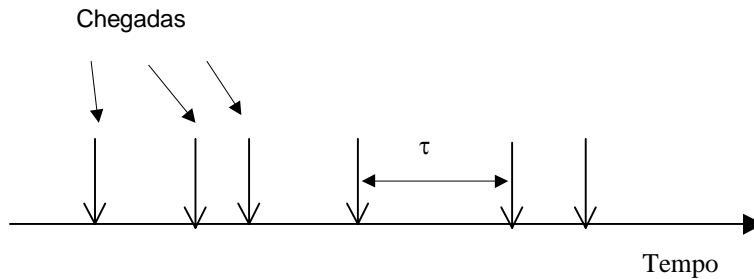



Figura 6.4 Instantes de chegada num processo de Poisson.

Considere-se agora um intervalo de tempo elevado e marque-se os instantes de chegada das chamadas (veja-se a Figura 6.4). O intervalo de tempo entre os instantes de chegada é representado por τ. Para um processo de chegada de Poisson τ é uma variável aleatória contínua com distribuição exponencial negativa, cuja função densidade de probabilidade é descrita por

$$p(\tau) = \lambda e^{-\lambda\tau} \qquad \tau \geq 0 \tag{6.9}$$

A partir de (6.9) é imediato o cálculo do valor médio de τ vindo

$$E(\tau) = \int_0^{\infty} \tau p(\tau) d\tau = \frac{1}{\lambda} \tag{6.10}$$

enquanto a variância vem dada por $\sigma_\tau^2 = 1/\lambda^2$. A probabilidade que o intervalo de tempo τ seja superior a *t* é dado por

$$P(\tau \geq t) = \lambda \int_t^{\infty} e^{-\lambda\tau} d\tau = e^{-\lambda t} \tag{6.11}$$

Exemplo 6.4
Considere uma central local com 20 000 assinantes. Tendo presente que cada assinante gera uma chamada por hora, determine a frequência com que ocorrem duas chamadas com um intervalo inferior a 0.02 s.

Solução: A probabilidade de o intervalo entre duas chamadas ser inferior a 0.02 s é dada por

$$P(\tau < 0.02) = 1 - P(\tau \geq 0.02) = 1 - e^{-\lambda x 0.02}$$

Por sua vez λ=20000x(1/3600)=5.56 chamadas/s, o que conduz a $P(\tau<0.02)=0.105$, ou seja, 10.5% das chamadas ocorrem com um intervalo inferior a 0.02 s. Como a taxa de chegada é de 5.56 chamadas por segundo obtém-se para a frequência pedida o valor de 5.56x0.105=0.584/s.

Outro factor importante na teoria do tráfego é a duração das chamadas T_c. Em algumas aplicações é suficiente conhecer o tempo médio de duração das chamadas, ou seja, $E(T_c)=h$. Em outros casos exige-se uma informação mais detalhada, como seja, o conhecimento da função densidade de probabilidade. A distribuição usada normalmente para caracterizar a duração das chamadas telefónicas é a distribuição exponencial negativa. Usando esta distribuição conclui-se que a probabilidade de T_c ser superior a *t* é dada por

$$P(T_c > t) = \exp(-t/h) \tag{6.12}$$

Para um grupo de *N* troncas o número de chamadas em progresso está compreendido entre 0 e *N*. O comportamento deste grupo pode-se descrever por uma cadeia de Markov com *N*+1 estados, como se representa na Figura 6.5.

O sistema está no estado *j* quando o número de troncas ocupadas é igual a *j*. A chegada de uma chamada faz aumentar o estado do sistema de 1, passando do estado *j* para o estado *k* (*k*=*j*+1). A terminação de uma chamada faz com que o estado do sistema decresça de 1, ou seja, se estiver no estado *k* passa para o estado *j*. *P*(*j*) é a probabilidade de o sistema se encontrar no estado *j* e *P*(*k*) a probabilidade do sistema se encontrar no estado seguinte *k*. $P_{j,k}$ é a probabilidade do sistema transitar do estado *j* para o estado *k*, enquanto $P_{k,j}$ é a probabilidade de transitar do estado *k* para o estado *j*. Assim, as probabilidades *P*(0), *P*(1),…*P*(*N*) designam-se por **probabilidades de estado**, enquanto $P_{j,k}$, $P_{k,j}$ são designadas por **probabilidades de transição.**

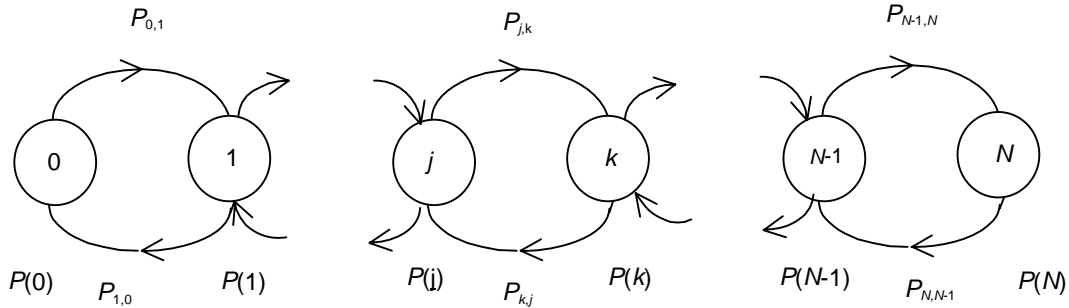



Figura 6.5 Diagrama de transição de estados para *N* troncas.

Considere-se um intervalo de tempo infinitesimal δt com início em t e admita-se que a probabilidade de ocorrerem dois ou mais eventos é desprezável. Os eventos que podem ocorrer em δt são os seguintes:

- Chegada de uma chamada, com probabilidade $P(a)$;
- Terminação de uma chamada, com probabilidade $P(b)$;
- Ausência de mudanças de estado, com probabilidade $1\text{-}P(a)\text{-}P(b)$.

A equação 6.3 permite concluir que o número médio de chamadas que chegam durante o intervalo de tempo δt é $A\delta t/h$. Atendendo ao facto de δt ser um intervalo infinitesimal, tem-se que $A\delta t/h<<1$ representa a probabilidade $P(a)$ de uma chamada chegar no intervalo de tempo δt. Assim,

$$P_{j,k} = P(a) = A\delta t / h \qquad (6.13)$$

Se o tempo médio de duração de uma chamada é h e o número de chamadas em progresso é k, espera-se que no intervalo h terminem em média k chamadas. O número médio de chamadas terminadas em δt será por conseguinte $k\delta t/h$. Com base num raciocínio idêntico ao do caso anterior pode-se escrever que

$$P_{k,j} = P(b) = k\delta t / h \qquad (6.14)$$

Tendo presente que a probabilidade de existirem j chamadas em progresso no instante t é $P(j)$, então a probabilidade de transição de j para k no intervalo de tempo δt é dada por

$$p(j \to k) = P(j)P(a) = P(j)A\delta t / h \qquad (6.15)$$

Se a probabilidade de haver k chamadas no instante t é $P(k)$, então a probabilidade de uma transição do estado k para o estado j durante δt é

$$p(k \to j) = P(k)P(b) = P(k)k\delta t / h \qquad (6.16)$$

Neste ponto convém fazer uma outra hipótese simplificativa, admitindo-se que existe um estado de **equilíbrio estatístico.** Segundo esta hipótese, o número médio de chamadas em progresso mantém-se constante (ou, haverá aproximadamente tantas chegadas ao sistema como terminações) o que faz com que a probabilidade de abandonar o estado *j* seja igual à probabilidade de mudança para ele, o que leva a escrever

$$p(j \to k) = p(k \to j) \qquad (6.17)$$

ou ainda

$$P(k) = \frac{A}{k} P(j) \qquad (6.18)$$

Repetindo a equação (6.18) *n* vezes chega-se a

$$P(n) = \frac{A^n}{n!} P(0) \qquad (6.19)$$

A hipótese de um tráfego puramente aleatório implica a existência de um número de fontes muito elevado. Assim, pode-se considerar que *n* varia entre zero e infinito, de modo que a condição de normalização de probabilidade adquire a forma

$$\sum_{n=0}^{\infty} P(n) = 1 \qquad (6.20)$$

Usando (6.19) e (6.20) é fácil demonstrar-se que

$$P(n) = \frac{A^n}{n!} e^{-A} \qquad (6.21)$$

A equação (6.21) mostra que o número de chamadas em progresso segue uma distribuição de Poisson, o que é consequência do facto de se admitir que a chegada de chamadas também é descrita por uma estatística de Poisson. Esta estatística requer um número de troncas infinito para escoar as chamadas. Como o número de troncas é necessariamente finito haverá um certo número de chamadas que são perdidas (ou atrasadas nas redes de pacotes) e a estatística das chamadas em progresso deixa de ser Poissoniana. Na secção seguinte ir-se-á analisar esta questão, considerando um sistema com perdas, isto é, um sistema em que as chamadas são perdidas no caso de não haver troncas livres.

6.1.6 Fórmula de Erlang para sistemas com perdas

Para se calcular o número de chamadas em progresso para um sistema com um número de troncas finito vai admitir-se que são válidas as seguintes condições:

- tráfego puramente aleatório;
- existe equilíbrio estatístico, isto é, o número de chamadas originadas num determinado período (ex. exemplo hora mais carregada) é em média igual ao número de chamadas terminadas nesse período;
- acesso completo, ou seja, se as chamadas que chegam são ligadas às troncas de saída por comutadores, estes não introduzem bloqueio (destaque-se que em grande número de casos práticos isto não é verdade, como se irá ver mais tarde);
- sistema com perdas, ou seja, as chamadas chegadas que não encontram troncas livres são perdidas.

A condição do sistema com perdas tem implícito, que qualquer chamada perdida não ocupa o equipamento durante nenhum tempo, e que o número máximo da chamadas em progresso não pode ultrapassar o número N de troncas disponíveis, isto é, $0 \le n \le N$. Nesta situação a equação (6.20) reescreve-se na forma

$$\sum_{n=0}^{N} P(n) = 1 \qquad (6.22)$$

o que, atendendo a (6.19), permite concluir que

$$P(0) = 1 \Big/ \sum_{n=0}^{N} \frac{A^n}{n!} \qquad (6.23)$$

Inserindo esta equação em (6.19) deduz-se que

$$P(n) = \frac{A^n / n!}{\sum_{k=0}^{N} A^k / k!} \qquad 0 \le n \le N \qquad (6.24)$$

Esta equação traduz a primeira distribuição de Erlang e representa a probabilidade de no conjunto das N troncas consideradas existirem n ocupadas (n chamadas em progresso).

A utilização da fórmula de Erlang vai permitir obter o grau de serviço. Para isso, admita-se que o tráfego A é oferecido sequencialmente às diferentes troncas. Assim, o tráfego A é, em primeiro lugar, todo oferecido à tronca 1. Nestas condições, a probabilidade de ocupação deste circuito é dada por

$$P(1) = \frac{A}{1+A} \qquad (6.25)$$

Durante a ocupação da tronca 1 o tráfego será desviado para a tronca 2, que vê oferecido o tráfego perdido pela tronca 1, ou seja

$$A_{of\,2} = A_{p1} = AP(1) = \frac{A^2}{1+A} \tag{6.26}$$

Do tráfego A é, portanto, transportado pela tronca 1

$$A_{t1} = A - A_{of\,2} = A(1-P(1)) \tag{6.27}$$

A probabilidade de ocupação da tronca 2, tendo presente que a tronca 1 está ocupada, será

$$P(2) = \frac{A^2/2}{1+A+A^2/2} \tag{6.28}$$

Deste modo, o tráfego oferecido à tronca 3, será o tráfego perdido pela tronca 2, ou

$$A_{of\,3} = A_{p2} = AP(2) = \frac{A^3/2}{1+A+A^2/2} \tag{6.29}$$

O tráfego transportado pela tronca 2, vem dado por

$$A_{t\,2} = A_{of\,2} - A_{of\,3} = A(P(1)-P(2)) \tag{6.30}$$

Idêntico raciocínio para a tronca N, permite escrever que

$$A_{ofN+1} = A_{pN} = AP(N) = \frac{A^{N+1}/N!}{\sum_{n=0}^{N} A^n/n!} \tag{6.31}$$

Tendo presente que existem unicamente N troncas, tem-se que o tráfego perdido pela tronca N (que teoricamente é igual ao tráfego oferecido à inexistente tronca N+1) dividido pelo tráfego inicial oferecido ao conjunto das N troncas, dá precisamente o grau de serviço oferecido por estas troncas, ou seja,

$$B = E_{1,N}(A) = \frac{A_{pN}}{A} = \frac{A^N/N!}{\sum_{n=0}^{N} A^n/n!} \tag{6.32}$$

Esta expressão é conhecida como fórmula de Erlang B, e desempenha um papel relevante na teoria do teletráfego. A fórmula de Erlang B pode-se ainda simplificar por

$$E_{1,N}(A) \approx \frac{A^N}{N!e^A} \tag{6.33}$$

expressão idêntica à da distribuição de Poisson, que é, por vezes, conhecida como fórmula de Grinstead. Essa simplificação baseou-se na aproximação

$$e^A \approx \sum_{i=0}^{N} \frac{A^i}{i!} \tag{6.34}$$

a qual só é válida para valores tais que $A^N / N! << 1$. A fórmula de Erlang B pode ainda ser calculada por um método iterativo recorrendo à fórmula de recorrência

$$E_{1,N}(A) = \frac{AE_{1,N-1}(A)}{N + AE_{1,N-1}(A)} \tag{6.35}$$

tendo presente que $E_{0,1}(A) = 1$. Outro procedimento usual para calcular a fórmula de Erlang faz uso de tabelas apropriadas (ver por exemplo a Tabela 6.1).

Exemplo 6.5
Considere-se um grupo de 50 troncas. Pede-se o tráfego oferecido que conduz a um grau de serviço de 0.001.

Solução: Para resolver esta questão faz-se uso da equação (6.33). Aplicando logaritmos a ambos os membros vem

$$\ln E_{1,N}(A) \approx N \ln A - \ln N! - A$$

Usando a fórmula de Stirling $N! = \sqrt{2\pi} N^{N+0.5} e^{-N}$

obtém-se

$$\ln E_{1,N}(A) = N \ln A - \ln \sqrt{2\pi} - (N + 0.5) \ln N + N - A$$

Como $E_{1,N}(A) = B = 0.001$ e N=50 conclui-se, usando a equação anterior, que A≈31.5E. Neste caso conclui-se facilmente que a condição $A^N / N! << 1$ não é verificada. Coloca-se a questão de saber qual a dimensão do erro cometido. Recorrendo à tabela 5.1 conclui-se que o valor exacto é A=32.5 E, ou seja o erro foi de 1 E.

6.1.7 Sistemas com filas de espera

A análise feita na secção anterior considerou um sistema com perdas. Neste tipo de sistema, qualquer chamada oferecida que não encontre troncas de saída livres é automaticamente perdida. Os sistemas de comutação analógicos passo-a-passo eram sistemas estritamente com perdas. Contudo, a introdução da comutação digital, veio criar

as condições para que as chamadas esperem em fila de espera, até que o equipamento, ou as troncas fiquem disponíveis para estabelecer a ligação.

Tabela 6.1 Tabela de tráfego

Número de troncas	Tráfego em Erlang			Número de troncas	Tráfego em Erlang		
	B=0.01	B=0.005	B=0.001		B=0.01	B=0.005	B=0.001
1	0.01	0.005	0.001	51	38.8	36.8	33.4
2	0.15	0.105	0.046	52	39.7	37.6	34.2
3	0.45	0.35	0.19	53	40.6	38.5	35.0
4	0.9	0.7	0.44	54	41.5	39.4	35.8
5	1.4	1.1	0.8	55	42.4	40.3	36.7
6	1.9	1.6	1.1	56	43.3	41.2	37.5
7	2.5	2.2	1.6	57	44.2	42.1	38.3
8	3.2	2.7	2.1	58	45.1	43.0	39.1
9	3.8	3.3	2.6	59	46.0	43.9	40.0
10	4.5	4.0	3.1	60	46.9	44.7	40.8
11	5.2	4.6	3.6	61	47.9	45.6	41.6
12	5.9	5.3	4.2	62	48.8	46.5	42.5
13	6.6	6.0	4.8	63	49.7	47.4	43.4
14	7.4	6.6	5.4	64	50.6	48.3	44.1
15	8.1	7.4	6.1	65	51.5	49.2	45.0
16	8.9	8.1	6.7	66	52.4	50.1	45.8
17	9.6	8.8	7.4	67	53.3	51.0	46.6
18	10.4	9.6	8.0	68	54.2	51.9	47.5
19	11.2	10.3	8.7	69	55.1	52.8	48.3
20	12.0	11.1	9.4	70	56.0	53.7	49.2
21	12.8	11.9	10.1	71	57.0	54.6	50.1
22	13.7	12.6	10.8	72	58.0	55.5	50.9
23	14.5	13.4	11.5	73	58.9	56.4	51.8
24	15.3	14.2	12.2	74	59.8	57.3	52.6
25	16.1	15.0	13.0	75	60.7	58.2	53.5
26	16.9	15.8	13.7	76	61.7	59.1	54.3
27	17.7	16.6	14.4	77	62.6	60.0	55.2
28	18.6	17.4	15.2	78	63.6	60.9	56.1
29	19.5	18.2	15.9	79	64.5	61.8	56.9
30	20.4	19.0	16.7	80	65.4	62.7	58.7
31	21.2	19.8	17.4	81	66.3	63.6	58.7
32	22.1	20.6	18.2	82	67.2	64.5	59.5
33	23.0	21.4	18.9	83	68.1	65.4	60.4
34	23.8	22.3	19.7	84	69.1	66.3	61.3
35	24.6	23.1	20.5	85	70.1	67.2	62.1
36	25.5	23.9	21.3	86	71.0	68.1	63.0
37	26.4	24.8	22.1	87	71.9	69.0	63.9
38	27.3	25.6	22.9	88	72.8	69.9	64.8
39	28.2	26.5	23.7	89	73.7	70.8	65.6
40	29.0	27.3	24.5	90	74.7	71.8	66.6
41	29.9	28.2	25.3	91	75.6	72.7	67.4
42	30.8	29.0	26.1	92	76.6	73.6	68.3
43	31.7	29.9	26.9	93	77.5	74.3	69.1
44	32.6	30.8	27.7	94	78.4	75.4	70.0
45	33.4	31.6	28.5	95	79.3	76.3	70.9
46	34.3	32.5	29.3	96	80.3	77.2	71.8
47	35.2	33.3	30.1	97	81.2	78.2	72.6
48	36.1	34.2	30.9	98	82.2	79.1	73.5
49	37.0	35.1	31.7	99	83.2	80.0	74.4
50	37.9	35.9	32.5	100	84.0	80.9	75.3

Neste caso, a chamada é atrasada em vez de ser perdida. Se o tráfego oferecido é maior que o número de troncas (designadas nos sistemas com filas de espera por servidores), a fila irá crescer continuamente em comprimento e a probabilidade de a chamada ser atrasada aproxima-se da unidade. Como há sempre a possibilidade de haver outras chamadas em fila de espera, a probabilidade de atraso é sempre superior à probabilidade de bloqueio para os sistemas com perdas, admitindo um sistema com as mesmas dimensões. A probabilidade de atraso D, pode-se relacionar com a probabilidade de bloqueio B, usando a segunda equação de Erlang, também designada por equação de Erlang C. Esta equação é dada por

$$E_{2,N}(A) = \frac{NE_{1,N}(A)}{N - A + AE_{1,N}(A)} \tag{6.36}$$

Onde $D = E_{2,N}(A)$ e $B = E_{1,N}(A)$. A validade da equação de Erlang C é baseada nos mesmos pressupostos que a equação de Erlang B. Além disso considera-se que:

- a fila pode aceitar um número infinito de chamadas em espera;
- as chamadas são processadas em função da ordem de chegada.

A primeira das condições é irrealizável. No entanto, desde que N seja superior a A a equação de Erlang C pode ser usada com um grau de precisão razoável.

6.1.8 Problemas

6.1 Durante uma hora de ponta foi observada a seguinte distribuição de chamadas telefónicas num grupo de 5 troncas correspondentes a 20 assinantes.

Tronca 1		Tronca 2		Tronca 3		Tronca 4		Tronca 5	
Início (min.)	Duração (min.)	Início (min.)	Duração (min.)	Início (min.)	Duração (min.)	Início (min.)	Duração (min.)	Início (min.)	Duração (min.)
8	4	2	4	6	6	4	10	2	4
22	6	10	6	22	2	20	6	12	10
38	2	24	2	38	6	30	4	36	6
48	8	34	8	50	8	46	12	48	2
--	--	52	4	--	-	-	--	54	2

Determine:

a) O andamento do tráfego instantâneo $K_i(t)$ ($K_i(t)$=1 se a tronca está activa; $K_i(t)$=0 se a tronca está inactiva).

b) O volume de tráfego e a intensidade de tráfego a partir deste.

c) A duração média das chamadas.

d) O tráfego médio por assinante.

6.2 Uma central telefónica serve 6000 assinantes. Durante uma hora de ponta a taxa de chamadas é de 1/3 de chamadas por assinante e a duração média das chamadas é de 3 minutos. Sabendo que o número médio de chamadas perdidas durante uma hora de ponta é de 10, calcule para esse intervalo:

a) O número médio de chamadas oferecidas à central.
b) O tráfego médio correspondente a estas chamadas.
c) O tráfego médio por assinante.
d) O grau de serviço (probabilidade de bloqueio).
e) O tráfego médio transportado.

6.3 Na hora mais carregada, um assinante com uma única linha, realiza quatro chamadas e recebe duas. A duração média de cada chamada é de 2 min. Qual é a probabilidade de que um assinante chamador encontre essa linha ocupada.

6.4) Considere que a um grupo de 5 troncas é oferecida uma chamada em cada 50 s e a duração média da chamada é de 100 s. Determine:

a) A intensidade do tráfego.
b) A probabilidade de bloqueio (grau de serviço).

6.5 A um grupo de 5 troncas é oferecido um tráfego de 2E. A duração média de uma chamada é de 3 minutos.

a) Qual é o número médio de chamadas oferecidas por hora?
b) Qual é a probabilidade de não serem oferecidas chamadas durante um período de tempo de 5 minutos?
c) Qual é o grau de serviço?
d) Admitindo que as diferentes troncas são testadas sequencialmente, qual é o tráfego transportado por cada tronca.
e) Qual é o tráfego perdido?

6.6 Durante a hora de ponta é oferecido, em média, um tráfego de 30 E a um grupo de troncas. Em média, o período total durante o qual todas as troncas estão ocupadas é de 12 segundos, sendo perdidas 2 chamadas. Determine o número médio de chamadas transportadas pelo grupo e a duração média das chamadas.

6.7 Numa central telefónica um grupo de 30 troncas proporciona uma probabilidade de bloqueio (grau de serviço) de 1%, quando é oferecido um tráfego de 20 E.

a) Obtenha uma fórmula de recorrência para a fórmula de Erlang-B.
b) Aplicando a fórmula deduzida determine a melhoria na probabilidade de bloqueio quando adiciona uma tronca.
c) Determine a deterioração da mesma probabilidade de bloqueio se se retirar uma tronca

6.8 Considere um número de assinantes M=1600 e um número de troncas de N=160. Para um grau de serviço de 0.01, determine o tráfego oferecido por assinante.

6.9 Considere uma central local para 8000 assinantes. Durante a hora de ponta cada assinante origina em média 0.2 chamadas e a duração média de cada chamada é de 3 minutos. Pretende-se conhecer o número de troncas de saída necessárias para garantir um grau de serviço de 0.005.

6.10 O espaçamento médio entre chamadas oferecidas a uma central telefónica na hora de ponta é de 3 segundos. Admitindo que a duração média das chamadas é de 4 minutos, determine:

a) A frequência com que ocorrem duas chamadas com um espaçamento inferior a 0.5 s.
b) O número de troncas necessárias para garantir um grau de serviço de 0.001.

6.11 Para um grau de serviço de 0.1% calcule a relação entre o tráfego oferecido e a eficiência do servidor, considerando como exemplo N=1, 2, 5, 10, 20, 50,100. Comente os resultados (tenha presente que a eficiência é definida por A/N).

6.1.9 Referências

[Be91] John Bellamy, *Digital Telephony*, John Wiley & Sons, 1991.

[Fl94] J.E. Flood, *Telecommunications Switching, Traffic and Networks*, Prentice Hall International, 1994.

[Sa79] Carlos Salema, *Estudo de Sistemas de Telecomunicações*, AEIST, 1979.

[Sc87] Misha Schwartz, *Telecommunication Networks*, Addison-Wesley, 1987.

6.2 Comutação

6.2.1 Aspectos genéricos

Em termos gerais o equipamento associado com qualquer central de comutação telefónica deverá realizar as seguintes funções:

- Sinalização
- Controlo
- Comutação

O papel básico da função de sinalização consiste em monitorizar a actividade das linhas telefónicas e enviar para o sistema de controlo informação sobre o estado destas, assim como, informação de controlo apropriada. O sistema de controlo por sua vez interage com o equipamento responsável pela comutação, no sentido deste estabelecer as ligações apropriadas, em resposta aos requisitos da sinalização. O sistema de comutação é responsável, por sua vez, por estabelecer e manter uma ligação física (circuito) entre os utilizadores envolvidos numa chamada, durante toda a duração da chamada, e ainda por interromper essa ligação logo que a chamada termine.

6.2.2 Evolução da comutação

As primeiras centrais de comutação usavam operadores humanos para ligar os circuitos de entradas aos circuitos de saída e daí designarem-se por centrais manuais. O desenvolvimento das redes telefónicos tornou impraticável este tipo de operação, conduzindo à automatização da comutação. As primeiras centrais telefónicas automáticas usavam comutação electromecânica. A evolução tecnológica conduziu em seguida às centrais electrónicas e posteriormente às centrais digitais[66].

Os tipos básicos de comutadores electromecânicos são o **comutador passo-a-passo** ou Strowger[67] e o **comutador de barras cruzadas** (*crossbar*). O elemento de comutação básico de um comutador passo-a-passo é o selector. Na sua forma mais simples, um selector consiste num conjunto de escovas, que se movem sobre um conjunto de contactos fixos, conhecido por banco de contactos, em resposta directa aos sinais decádicos gerados pelo telefone. O movimento das escovas pode ocorrer, quer na vertical, quer na horizontal. O primeiro dígito do sinal decádico faz com que o primeiro selector se mova na vertical um número de níveis correspondente ao dígito marcado, e em seguida na horizontal até encontrar uma linha de saída livre para o segundo selector. O segundo selector responde ao segundo dígito do sinal de endereçamento marcado de modo similar. O processo contínua através de tantos selectores quantos os exigidos pela dimensão do comutador.

[66] A primeira central digital foi introduzida em Portugal em 1986.
[67] As últimas centrais analógicas baseadas na comutação Strowger só foram retiradas da rede telefónica pública em Portugal em meados dos anos 90.

Um sistema passo-a-passo usa um método de controlo designado por **controlo directo progressivo**. Com este método, cada andar do comutador responde directamente aos dígitos marcados pelo assinante chamador, sendo a conexão estabelecida progressivamente através do comutador. Com este método de controlo, os elementos de controlo do comutador são distribuídos ao longo do sistema na forma de relés que controlam individualmente cada selector.

Em alternativa, o comutador de barras cruzadas usa **controlo centralizado** para estabelecer as ligações. Neste caso, o sistema de controlo começa por armazenar os dígitos marcados, para os processar em seguida e gerar os sinais apropriados para controlar a matriz de comutação no estabelecimento de conexões.

No comutador de barras cruzadas, a matriz de comutação tem realmente a aparência de uma matriz constituída por barras horizontais e verticais. Estas barras são operadas por electroímanes activados pela corrente gerada pelo sistema de controlo, e os pontos de cruzamento são estabelecidos através do contacto mecânico entre essas barras.

A comutação electrónica constituiu uma nova etapa na evolução das tecnologias de comutação. Nas centrais de comutação baseadas nesta tecnologia os pontos de cruzamento eram estabelecidos usando os relés de palheta (*reed*). O relé de palheta é constituído por um par de contactos, feitos de um material magnético, que são inseridos dentro de uma cápsula de vidro onde se faz vácuo. A cápsula é colocada ao longo do eixo de uma bobina, de modo que, quando esta é excitada gera um campo magnético que contribui para atrair os dois contactos e activar o ponto de cruzamento. Outra particularidade das centrais electrónicas residia no facto das funções de controlo serem realizadas por computador, designando-se por isso centrais SPC (*stored program control).* O computador tem a vantagem de poder executar diferentes funções com o mesmo hardware, fazendo correr diferentes programas, o que não era possível de ser realizado nos comutadores anteriores.

Todas estas técnicas de comutação são técnicas de **comutação espacial.** Um comutador espacial liga e desliga contactos físicos usando uma matriz de pontos de cruzamento. A qualquer conexão estabelecida através de um comutador espacial corresponde um caminho eléctrico permanente (circuito), que se mantém durante toda a duração da chamada.

A comutação digital introduz uma outra dimensão na comutação: o tempo. Por isso, os comutadores digitais, para além da comutação espacial, usam também **comutação temporal**. A necessidade de comutação temporal compreende-se facilmente se se tiver presente que os sistemas de linha presentes na entrada da central não são circuitos individuais, mas sim sistemas a 2 Mbit/s (32 canais), ou 1.5 Mbit/s (24 canais). As centrais de comutação digital são também centrais SPC, ou seja, centrais controladas por software.

6.2.3 Comutação espacial

A Figura 6.6 mostra uma matriz de comutação espacial constituída por N linhas de entrada e M linhas de saída, apresentando NM pontos de cruzamento. Os pontos de cruzamento podem ser implementados usando contactos mecânicos (comutador *crossbar*), relés de palheta (comutador electrónico) ou portas lógicas (comutador digital).

Essa matriz não apresenta bloqueio de interligação (conectividade total): não há chamadas perdidas a não ser que os circuitos de saída estejam congestionados. O número de conexões que podem ser feitas é M (se $M>N$) ou N (se $N<M$).

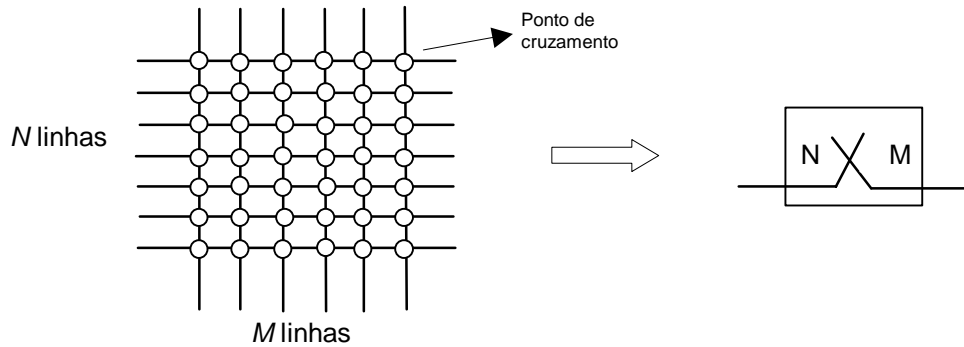



Figura 6.6 Matriz de comutação com uma arquitectura de barras cruzadas.

Um parâmetro característico do desempenho das matrizes de comutação é a sua complexidade, que está directamente associada ao seu custo. Tradicionalmente, a complexidade de uma matriz era dada pelo número de pontos de cruzamento. Hoje em dia, nos comutadores digitais existem outros parâmetros que contribuem para o custo da matriz, como sejam o número de portas lógicas por *chip,* o número de chips por *board,* o número de pontos de cruzamento por *board*, etc. Contudo, como é complexo caracterizar o custo de uma matriz entrando com todos esses factores, vai-se considerar neste estudo, que o custo da matriz (complexidade) é determinada fundamentalmente pelo número de pontos de cruzamento. Considerando uma matriz de barras cruzadas quadrada (N=M), a sua complexidade é dada por

$$C(1) = N^2 \tag{6.37}$$

ou seja, aumenta com o quadrado da dimensão da matriz. Contudo, a eficiência (indicada pela percentagem de pontos de cruzamento que podem estar simultaneamente activos, isto é, $N/N^2=1/N$) decresce inversamente com N. Por exemplo, um comutador com 100 entradas e 100 saídas requer 10000 pontos de cruzamentos, dos quais só 1% podem estar simultaneamente activos. Outra desvantagem da estrutura referida é a sua baixa fiabilidade, resultante do facto de qualquer avaria num ponto de cruzamento fazer com que haja uma ligação que não se pode efectuar. É, pois, desejável encontrar uma arquitectura alternativa, que garanta ainda conectividade total, mas que: 1) permita caminhos alternativos (de modo que a fiabilidade do sistema aumente); 2) partilhe pontos de cruzamento entre as diversas ligações possíveis (de modo a aumentar a eficiência).

Estes objectivos são atingidos usando **arquitecturas multiandar**, onde cada andar é baseado em sub-matrizes com conectividade total (sem bloqueio). Outra vantagem adicional destas arquitecturas reside na possibilidade de combinar diferentes tecnologias para realizar uma determinada matriz. Esta facilidade é particularmente útil na comutação

digital, onde é usual combinar andares baseados em comutação temporal com andares baseados em comutação espacial. Na Figura 6.7 representa-se uma matriz espacial com N linhas de entrada e N linhas de saída usando dois andares. As N linhas de entrada são divididas por $k=N/n$ grupos de n linhas. Cada um destes grupos é ligado a uma sub-matriz $n\times k$. Tem-se, assim, que o andar de entrada é constituído por k sub-matrizes $n\times k$, enquanto o andar de saída é constituído por k sub-matrizes $k\times n$.

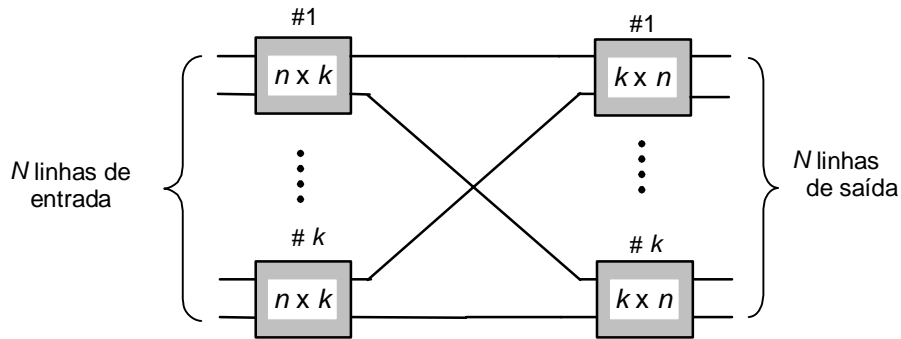



Figura 6.7 Matriz de comutação espacial NxN com dois andares.

Para ajuizar da possível melhoria da eficiência associada a esta arquitectura é necessário calcular o número de pontos de cruzamento, o qual é dado por

$$C(2) = 2nk^2 = 2N^2 / n \tag{6.38}$$

Outro parâmetro importante na arquitectura em análise é o número total de ligações existentes entre as sub-matrizes do primeiro andar e as do segundo andar. Este número é igual a N^2/n^2. Deste modo, o número de pontos de cruzamento varia com $1/n$, enquanto o número de ligações varia com $1/n^2$. Se para reduzir o número de pontos de cruzamento o valor de n é feito muito elevado, então o número de ligações interiores reduz-se. Assuma-se que o número de ligações interiores é igual ao número de circuitos de entrada e circuitos de saída ($k^2=N$). Nesta situação, o número de pontos de cruzamento vem dado por

$$C(2) = 2N^{3/2} \tag{6.39}$$

Comparando a equação (6.37) e (6.38) verifica-se que com a utilização de dois andares se conseguiu uma redução do custo da matriz. Efectivamente, para o exemplo apresentado anteriormente necessitava-se neste caso de 2 000 pontos de cruzamento, em vez dos 10 000 requeridos pela arquitectura monoandar.

Embora a utilização dos dois andares tenha conduzido a um aumento da eficiência, verifica-se, contudo, que esta estrutura apresenta alguns problemas, devido ao facto de só ser possível estabelecer uma ligação entre as sub-matrizes de andares diferentes. Em primeiro lugar apresenta bloqueio, em segundo apresenta uma fiabilidade baixa. Para ultrapassar estas limitações pode-se recorrer a matrizes com três ou mais andares.

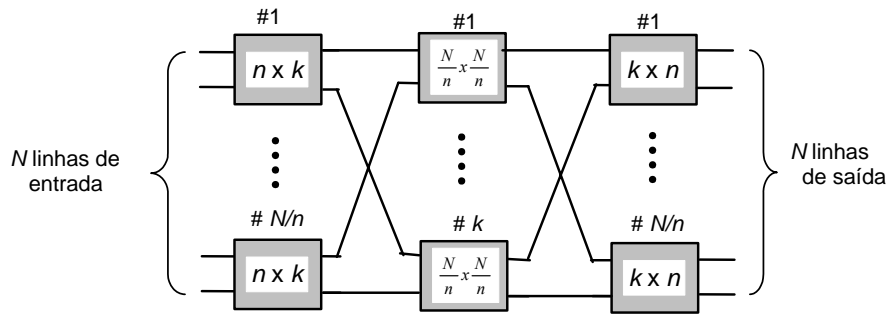



Figura 6.8 Matriz de comutação espacial *N*x*N* com três andares.

Na Figura 6.8 representa-se uma matriz com três andares. A principal especificidade desta estrutura reside no facto de existirem k ligações entre cada par de sub-matrizes no primeiro e no último andar, permitindo aumentar a fiabilidade e eliminar o bloqueio através de uma escolha adequada do valor de k. É imediato verificar que o número total de pontos de cruzamento neste caso é dado por

$$C(3) = 2kN + k\left(\frac{N}{n}\right)^2 \tag{6.40}$$

Este número pode ser minimizado, depois de se estabelecer uma relação entre n e k, como se irá verificar na secção seguinte.

6.2.3.1 Matrizes multiandar sem bloqueio de interligação

Uma matriz apresenta bloqueio de interligação nulo, quando é sempre possível estabelecer uma ligação entre uma entrada livre e uma saída livre, independentemente do estado da rede de interligação no instante de estabelecimento dessa ligação. As matrizes sem bloqueio podem-se dividir em três tipos:

- Sem bloqueio em sentido estrito;
- Sem bloqueio em sentido lato;
- Sem bloqueio rearranjáveis.

Numa estrutura sem bloqueio em sentido estrito é sempre possível ligar uma entrada livre a uma saída arbitrária livre, independentemente do conjunto de interligações já estabelecidas e do algoritmo de conexão usado. Uma matriz sem bloqueio em sentido lato garante que uma ligação entre uma entrada livre e uma saída livre nunca é bloqueada, desde que se use um algoritmo adequado para estabelecer as ligações. Finalmente, nas matrizes rearranjáveis pode haver necessidade de rearranjar as interligações de modo a garantir que seja sempre possível ligar uma entrada livre a qualquer saída livre. A Figura 6.9 mostra um exemplo de uma matriz rearranjável, designada por matriz de Benes. Com

o tipo de interligações apresentadas na Figura 6.9 (a) não é possível estabelecer algumas conexões devido à existência de bloqueio. Porém, através do conveniente rearranjo dessas interligações é possível ultrapassar a situação de bloqueio, como se conclui da Figura 6.9 (b).

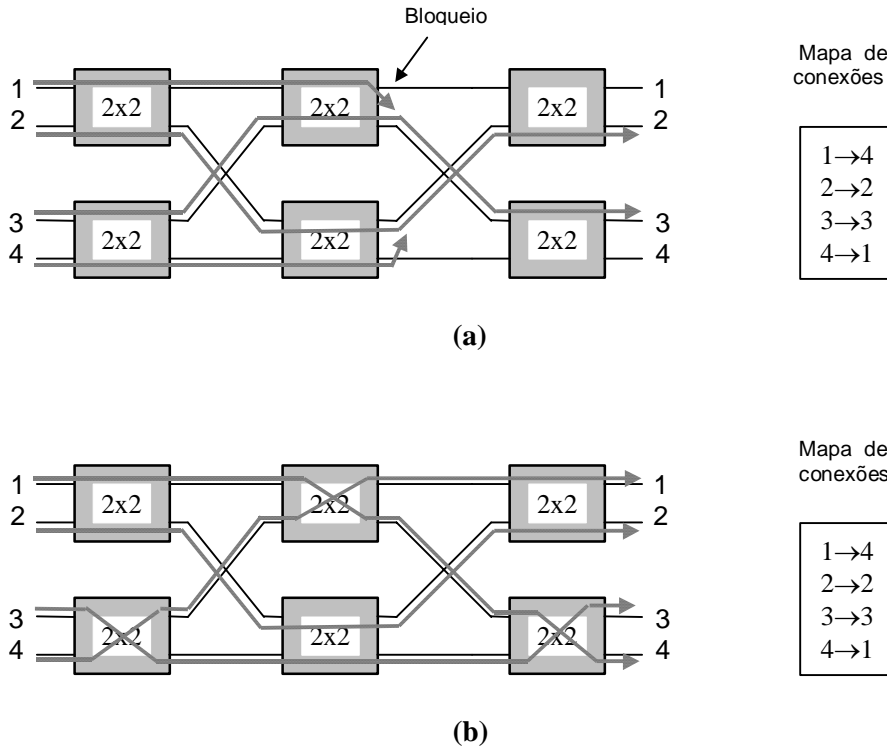



Figura 6.9 Matriz de comutação 4x4 com arquitectura de Benes. (a) Com bloqueio; (b) Eliminou-se o bloqueio depois de rearranjar as interligações.

Para que uma matriz multiandar, usando sub-matrizes com n entradas no andar de entrada e sub-matrizes com n saídas no andar de saída, não introduza bloqueio é necessário que as ligações intermédias permitam n conexões simultâneas entre cada sub-matriz de entrada e cada sub-matriz de saída. Deste modo, para que uma matriz rearranjável com três andares não introduza bloqueio é necessário verificar-se a condição [Be62]

$$k \geq n \tag{6.41}$$

Usando as equações (6.40) e (6.41) tira-se que o número mínimo de pontos de cruzamento para as matrizes rearranjáveis com três andares é dado por

$$C(3) = N(2n + N/n) \tag{6.42}$$

O número mínimo de pontos de cruzamento é determinado facilmente, tratando n como uma variável contínua, diferenciando C em ordem a n e igualando a zero, obtendo-se

$$n = \sqrt{N/2} \tag{6.43}$$

o que conduz a

$$C(3) = 2\sqrt{2}N^{3/2} \tag{6.44}$$

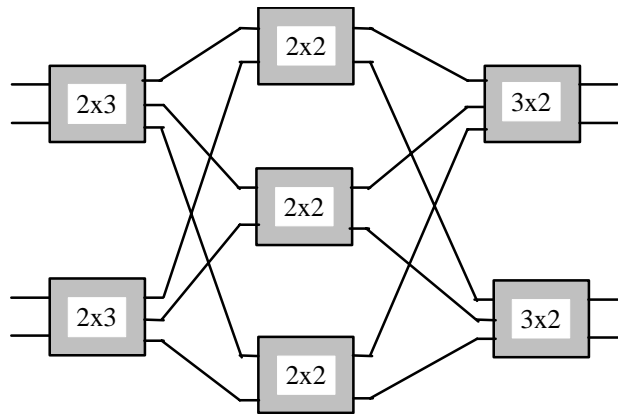



Figura 6.10 Matriz de comutação 4×4 com arquitectura de Clos (sem bloqueio em sentido estrito).

Na Figura 6.10 representa-se uma matriz de comutação 4×4 sem bloqueio em sentido estrito. O número de sub-matrizes do andar intermédio é obtido pelo teorema de Clos [Cl53] que se apresenta a seguir.

Teorema de Clos: Uma matriz de comutação com três andares (Figura 6.8) não introduz bloqueio de interligação em sentido estrito se e só se verificar a condição

$$k \geq 2n - 1 \tag{6.45}$$

Este teorema pode ser demonstrado considerando duas sub-matrizes, uma pertencente ao andar de entrada (sub-matriz A) e a outra ao de saída (sub-matriz C) de uma matriz com três andares (ver Figura 6.11), com um grau de ocupação tal que só é possível o estabelecimento de mais uma ligação entre a entrada i e a saída j. Assim, das n entradas da sub-matriz A n-1 estão ocupadas, e das n saídas da sub-matriz C também n-1 estão ocupadas. A pior situação, no que diz respeito ao bloqueio, ocorre quando as ligações entre as n-1 entradas da sub-matriz A e o andar de saída são feitas através de diferentes sub-matrizes do andar intermédio, que por sua vez são diferentes das sub-matrizes usadas para estabelecer as ligações entre o andar de entrada e as n-1 saídas da sub-matriz C. Isto quer dizer que, não há nenhuma ligação estabelecida entre a sub-matriz A e a sub-matriz C. Como é necessária uma sub-matriz adicional no andar intermédio são necessárias no mínimo $(n-1)+(n-1)+1 = 2n-1$ sub-matrizes no segundo andar para obter uma matriz de três andares sem bloqueio em sentido estrito.

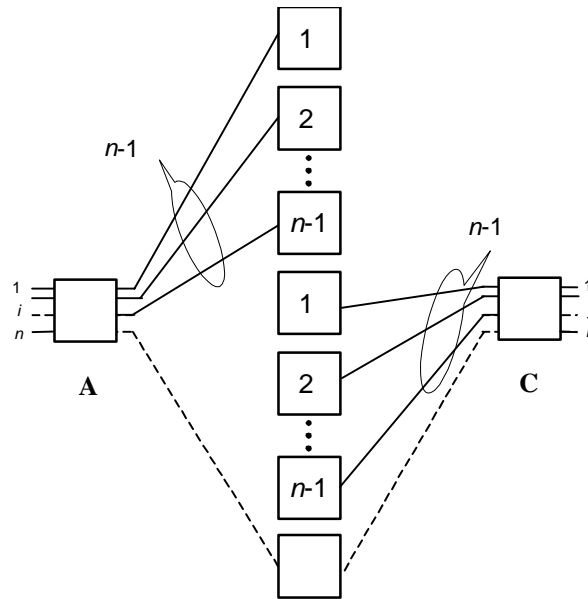



Figura 6.11 Princípio usado na demonstração do teorema de Clos.

Usando o valor mínimo de k que verifica o teorema de Clos, pode obter-se a partir de (6.40) o número de pontos de cruzamento para uma matriz de três andares sem bloqueio em sentido estrito, vindo

$$C(3) = 2(2n-1)N + (2n-1)\left(\frac{N}{n}\right)^2 \tag{6.46}$$

Esta expressão mostra que, para uma dada matriz NxN o número de pontos de cruzamento necessários depende do número de entradas n das sub-matrizes do andar de entrada (ou das saídas das sub-matrizes do andar de saída visto que se está a admitir uma estrutura simétrica). A dimensão óptima dessas sub-matrizes é determinada facilmente diferenciando $C(3)$ em ordem a n e igualando a zero, obtendo-se a seguinte equação:

$$2n^3 - nN + N = 0 \tag{6.47}$$

Para valores elevados de n a solução desta equação é aproximada por

$$n \cong \sqrt{N/2} \tag{6.48}$$

Substituindo este valor de n na expressão do número de pontos de cruzamento $C(3)$, obtém-se, finalmente, para o número mínimo de pontos de cruzamento, para arquitecturas de três andares sem bloqueio em sentido estrito, a seguinte expressão:

$$C(3) = 4N(\sqrt{2N} - 1) \tag{6.49}$$

Como se pode verificar, para valores grandes de *N* (*N*>25) a expressão anterior conduz a valores inferiores de *C*(1), sendo a diferença tanto maior quanto maior for *N*. O valor de *C* ainda pode ser reduzido usando arquitecturas com mais de três andares. As matrizes com vários andares baseadas na arquitectura de Clos podem ser construídas recursivamente a partir de matrizes com um número inferior de andares. Por exemplo, uma matriz de cinco andares (ver Figura. 6.12) pode ser vista como uma estrutura com três andares, na qual as sub-matrizes do andar intermédio são substituídas por matrizes de três andares sem bloqueio [Cl53]. O número de estruturas de três andares que é requerido será, por conseguinte, igual a 2*n*-1.

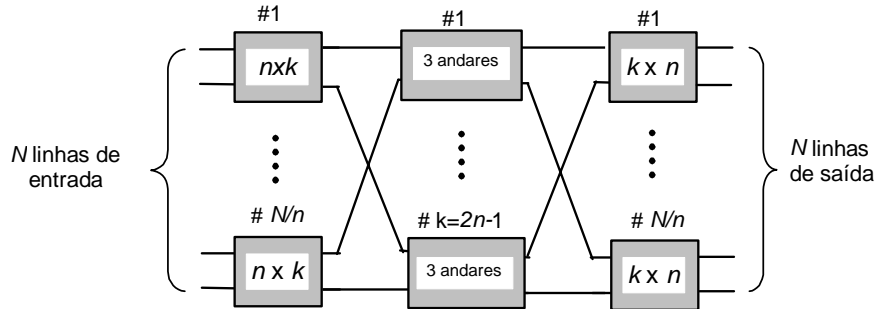



Figura 6.12 Matriz espacial com arquitectura de Clos de cinco andares.

Pode-se demonstrar que o número de pontos de cruzamento é mínimo quando $n = (2N)^{1/3}$ e é dado por

$$C(5) = 12 \times 2^{1/3} N^{4/3} - 14N + 2^{5/3} N^{2/3} \qquad (6.50)$$

De modo idêntico, uma matriz com sete andares pode ser desenhada partindo de uma matriz com três andares, na qual as sub-matrizes do segundo andar são substituídas por estruturas com cinco andares [Fl95].

Normalmente, as matrizes sem bloqueio em sentido estrito não são muito usadas em sistemas de comutação espaciais, porque não são económicas. Deste modo, no projecto desses sistemas é necessário contabilizar o efeito do bloqueio de interligação, o que irá ser feito na secção seguinte.

6.2.3.2 Cálculo da probabilidade de bloqueio de interligação através dos grafos de Lee

A probabilidade de bloqueio de interligação em matrizes de comutação espacial pode ser realizada usando um formalismo desenvolvido por Lee em 1955 [Le58]. Este formalismo faz uso da teoria dos grafos para representar a matriz. Um grafo é uma colecção de pontos designados por **vértices** ligados por linhas designadas por **arcos**.

Na Figura 6.13 representa-se o grafo de uma matriz espacial 3x3. Como se pode ver esse grafo baseia-se na representação das interligações entre as sub-matrizes, que são simbolizadas através de pontos. A representação de uma matriz através de um grafo pode

ainda ser simplificada desenhando unicamente os caminhos usados para estabelecer uma ligação entre uma determinada entrada e uma determinada saída. O grafo correspondente designa-se por **grafo de canal** (ver Figura 6.13). A primeira fase da aplicação da teoria de Lee consiste em obter o grafo de canal da matriz de comutação em análise.

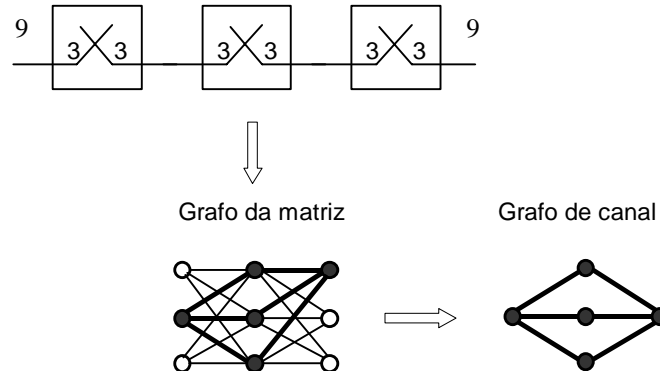



Figura 6.13 Representação de uma matriz espacial 3x3 através de um grafo.

No cálculo da probabilidade de bloqueio representa-se por p a fracção de tempo que uma determinada ligação está ocupada, ou seja, a probabilidade de a ligação estar ocupada. A probabilidade de ocupação p (ou simplesmente ocupação) em cada andar é obtida dividindo o tráfego total oferecido pelo número de ligações presentes nesse andar [Fl95].

O método de Lee admite que a ocupação simultânea de diferentes ligações constitui eventos estatisticamente independentes. De acordo com essa hipótese tem-se que:

- Se uma conexão entre dois nós é constituída por n caminhos alternativos, cada um com ocupação p, então a probabilidade dessa conexão estar bloqueada é dada por

$$B = p^n \tag{6.51}$$

- Se um caminho é constituído por n ligações em série, cada uma com a probabilidade q=1-p de estar livre, então a probabilidade de bloqueio é dada por

$$B = 1 - q^n \tag{6.52}$$

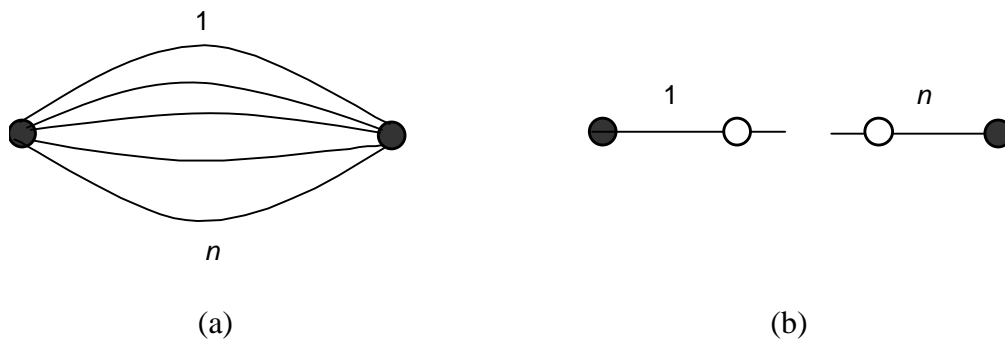



Figura 6.14 (a) Caminhos alternativos; **(b)** Caminho constituído por n ligações em série.

Como qualquer grafo de canal resulta da associação de caminhos alternativos e ligações em série, a probabilidade de bloqueio para qualquer estrutura de comutação é calculada através da utilização repetida daquelas duas equações. Exemplifica-se, em seguida, a aplicação deste método ao cálculo da probabilidade de bloqueio de uma matriz de comutação com três andares, cujo grafo de canal é apresentado na Figura 6.15. Note-se que, existem k caminhos alternativos, correspondendo cada um deles a uma das sub-matrizes de comutação do andar intermédio, e cada um destes caminhos envolve duas ligações. Assim, designando por p' a probabilidade de qualquer ligação entre andares estar ocupada, a probabilidade de bloqueio de interligação da estrutura pode ser calculada do seguinte modo:

$$
\begin{aligned}
B &= \text{Probabilidade dos } k \text{ caminhos estarem ocupados} \\
&= (\text{Probabilidade de um caminho qualquer estar ocupado})^k \\
&= (\text{Probabilidade de que pelo menos uma ligação no caminho esteja ocupada})^k \\
&= (1 - (q')^2)^k
\end{aligned}
\tag{6.53}
$$

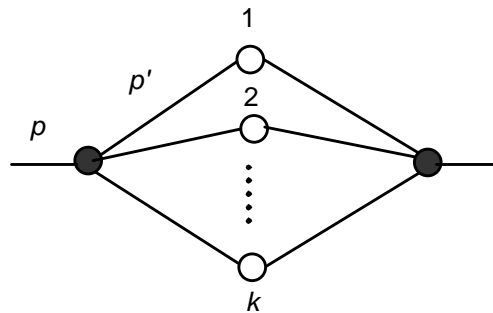



Figura 6.15 Grafo de canal de uma matriz com três andares.

onde $q' = 1 - p'$ é a probabilidade de uma ligação entre andares se encontrar livre. Admita-se que a probabilidade de ocupação das linhas de entrada da matriz p é conhecida. Então, a probabilidade de uma ligação entre andares estar ocupada é dada por

$$p' = p\frac{n}{k} \tag{6.54}$$

onde n é o número de entradas da sub-matriz do andar de entrada (ou saídas do andar de saída), usadas para estabelecer a conexão. Esta equação resulta do facto de um certo número de entradas (ou saídas) ocupadas implicar a existência de um número idêntico de saídas do primeiro andar (ou entradas do terceiro andar) ocupadas. Por conseguinte, se houver um número de ligações entre andares $\beta = k/n$ vezes superior ao número de entradas (saídas), será de esperar que a percentagem de ligações ocupadas seja reduzida de β.

No raciocínio apresentado está implícito que k é maior que n, o que implica que o primeiro andar da matriz proporciona expansão do tráfego oferecido à matriz. Pode, no entanto, também acontecer o inverso, ou seja $\beta < 1$ (k menor que n) implicando uma operação de concentração. Normalmente, nas centrais locais e em alguns PPCA de maiores dimensões usa-se concentração nos andares de entrada das matrizes de comutação, já que as linhas de acesso são fracamente utilizadas (percentagem de utilização de 5 a 10 %). Nas centrais tandem e de trânsito como as troncas são fortemente utilizadas é usual recorrer-se a expansão no primeiro andar para assegurar níveis de bloqueio de interligação reduzidos.

Atendendo a (6.53) e (6.54) obtém-se a seguinte expressão para a probabilidade de bloqueio de uma matriz de comutação com três andares:

$$B = \left(1 - (1 - pn/k)^2\right)^k \tag{6.55}$$

Com base nesta expressão é possível obter uma configuração para a matriz (valores de n e k) que permita garantir uma determinada probabilidade de bloqueio.

Tabela 6. 2. Número de pontos de cruzamento para matrizes com três andares com bloqueio e sem bloqueio

Dimensão da matriz	n	k	Factor de concentração	Número de pontos de cruzamento	
				Com bloqueio (0.01)	Sem bloqueio
32	4	3	1.333	384	896
128	8	4	2	2 048	7 680
512	16	6	2.667	12 288	63 488
2 048	32	9	3.556	73 728	516 096
8 192	64	14	4.571	458 522	$\approx 4.2\times10^6$
32 768	128	22	5.818	$\approx 2.9\times10^6$	$\approx 33\times10^6$
131 072	256	39	6.564	$\approx 20.4\times10^6$	$\approx 268\times10^6$

Na Tabela 6.2 representa-se o número de pontos de cruzamento para uma matriz com três andares considerando uma situação com bloqueio e uma outra sem bloqueio. Para a primeira situação admitiu-se uma probabilidade de bloqueio da ordem de 0.01 e uma probabilidade de ocupação das linhas de entrada de 0.1. Constata-se, facilmente, a significativa redução do número de pontos de cruzamento, conseguida à custa da introdução de bloqueio e de concentração. Nos casos em que não é possível introduzir concentração (probabilidade de ocupação elevada), será de esperar que a redução dos pontos de cruzamento conseguida com bloqueio seja mais modesta e será, porventura, necessário recorrer a estruturas com mais de três andares para conseguir reduções significativas do número de pontos de cruzamento.

6.2.4 Comutação digital

A comutação digital também é conhecida por **comutação por divisão no tempo**, essencialmente porque neste tipo de comutação o circuito associado a uma conversação só é activado durante o *time-slot* correspondente a essa conversação. Assim, pode-se afirmar que a comutação digital consiste em ligar um determinado *time-slot* de uma determinada entrada do comutador a um certo *time-slot* de uma certa saída do comutador. Este tipo de operação implica necessariamente a conversão dos sinais de entrada para um formato TDM, antes que qualquer operação de comutação tenha lugar.

Na figura 6.16 mostra-se um exemplo de um sistema de comutação digital em que o canal 3 (*time-slot* 3) da linha de entrada 1 é comutado para o canal 31 (*time-slot 31*) da linha de saída *N.* Assume-se que os sinais TDM têm 32 *time-slots.* Como se pode ver, a comutação tem lugar, quer no domínio do espaço (da linha 1 para a linha *N*), quer no domínio do tempo (do *time-slot* 3 para o *time-slot* 31). A exigência de uma comunicação bidireccional é aqui garantida comutando o canal 31 da entrada *N* para o canal 3 da saída 1, o qual vai ser usado para transmitir a informação em sentido inverso.

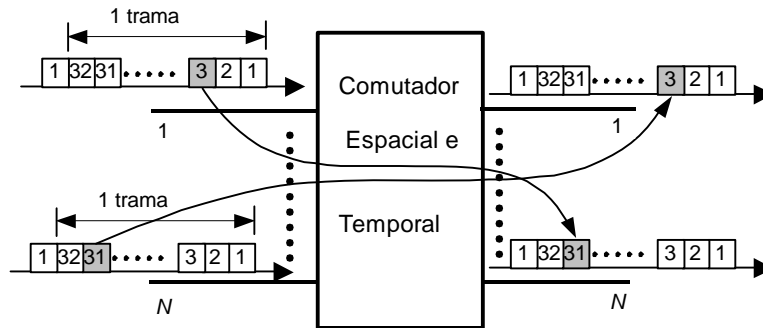



Figura 6.16 Princípio de um sistema de comutação digital.

Como se exemplificou, a comutação digital envolve operações de comutação espacial e temporal. Nas secções seguintes irão ser descritos os blocos responsáveis por essas operações antes de serem combinados para construir sistemas de comutação digital com vários andares.

6.2.4.1 Comutador digital espacial

Um comutador digital espacial pode ser visto como uma matriz de pontos de cruzamento, cuja configuração (estado dos pontos de cruzamento) pode ser alterada de *time-slot* para *time-slot.* Esses pontos de cruzamento são implementados usando portas lógicas rápidas, as quais são controladas por memórias designadas aqui por **memórias de conexão**.

A Figura 6.17 ilustra o funcionamento de um comutador digital espacial [Re95]. O comutador compreende uma matriz $n \times m$ e um conjunto de memórias de conexão. As n linhas da matriz correspondem às vias PCM/TDM de entrada (cada uma com w *time-slots*), enquanto as m colunas correspondem às vias PCM/TDM de saída (com o mesmo número w de *time-slots*).

Os pontos de cruzamento em cada uma das *m* colunas são controlados por uma memória de conexão, com capacidade para armazenar um número de palavras igual ao número de *time-slots w.* A cada ponto de cruzamento numa coluna é atribuído um endereço binário único. O endereço apropriado é, então, usado para activar o ponto de cruzamento adequado para estabelecer uma ligação entre uma determinada via de entrada e uma determinada via de saída. Os endereços seleccionados são armazenados na memória de conexão, ordenados por *time-slot*, de acordo com o plano de interligações. Isto é, para a coluna 1, o endereço do ponto de cruzamento a ser activado durante o *time-slot* 1 é armazenado na primeira célula da memória de conexão da coluna 1, o endereço do ponto de cruzamento a ser activado no *time-slot* 2 é armazenado na célula 2 e assim sucessivamente. A dimensão das células (palavras) da memória de endereços deverá ser suficiente para armazenar um endereço binário para os *n* pontos de cruzamento de cada coluna, mais um endereço para manter todos os pontos de cruzamento desactivados. São necessários (n+1) endereços, cada um identificado por uma palavra binária de comprimento $\log_2(n+1)$.

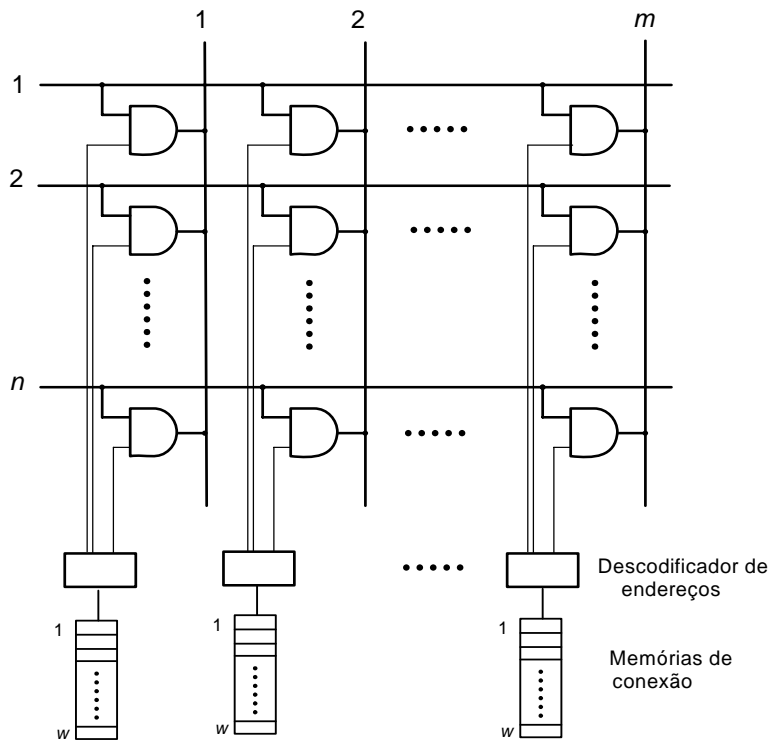



Figura 6.17 Estrutura de uma matriz de comutação digital espacial.

A acção do sistema de controlo consiste, em primeiro lugar, em carregar os endereços dos pontos de cruzamento de acordo com o plano de interligações, para em seguida proceder à leitura do conteúdo de cada célula de memória durante o *time-slot* apropriado. O endereço lido é usado para seleccionar o ponto de cruzamento correspondente, o qual permanecerá activado durante todo o *time-slot.* Este processo continua até que todas as células de memória tenham sido lidas e os correspondentes pontos de cruzamento processados. O processo de leitura de células e selecção dos pontos

de cruzamento repete-se com um período igual à duração da trama, durante o qual uma palavra (8 bits) de cada uma das vias PCM/TDM de entrada deverá ser permutada para uma via PCM/TDM de saída. Novas chamadas, ou a terminação das chamadas existentes, são efectuadas, respectivamente, através da inserção ou apagamento dos endereços dos pontos de cruzamento nas células apropriadas.

O facto de as linhas de entrada e dos pontos de cruzamento serem partilhados no domínio do tempo leva a uma redução significativa do número de pontos de cruzamento, em comparação com as matrizes espaciais analógicas. Por exemplo, considerando unicamente uma direcção de transmissão, a interligação de 1600 canais de entrada com 1600 canais de saída requer unicamente uma matriz com $50 \times 50 = 2500$ pontos de cruzamento, desde que se usem vias PCM/TDM com 32 *time-slots*. Por sua vez, uma matriz espacial analógica requereria $1600 \times 1600 = 2.56 \times 10^6$ pontos de cruzamento, considerando de novo uma única direcção de transmissão.

6.2.4.2 Comutador digital temporal

Como já se referiu, a comutação temporal consiste em transferir o conteúdo de um *time-slot,* para um outro *time-slot* não coincidente com o primeiro. Esta operação é realizada por um sistema de intercâmbio de *time-slots* designado por TSI (*Time-Slot Interchange*), cuja acção consiste em atrasar de modo apropriado os diferentes *time-slots* presentes na sua entrada. Assim, por exemplo, a permuta do TS2 (*time slot* 2) da via de entrada de um TSI para o TS7 da via de saída implica um atraso de 5 *time-slots.*

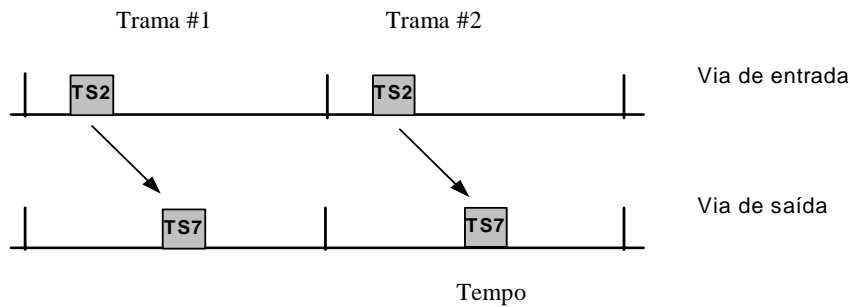
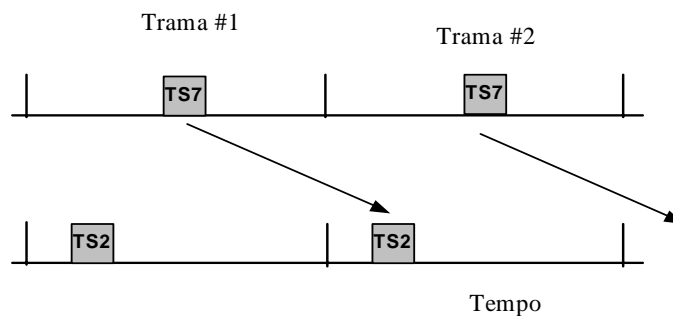



Figura 6.18 Princípios da comutação digital temporal.

A Figura 6.18a ilustra exactamente o processo de permuta desses *time-slots*, considerando duas tramas consecutivas. Como o TS7 aparece depois do TS2, o conteúdo dos canais de entrada e saída é mantido na mesma trama. Contudo, se a comutação temporal ocorresse entre um *time slot* de entrada e um *time slot* de saída, cuja ordem na trama fosse anterior à do primeiro, o atraso requerido implicava que o conteúdo do primeiro *time-slot* deveria ser atrasado até à trama seguinte. A Figura 6.18b ilustra o caso da comutação temporal do TS7 para o TS2. Aqui, o canal de entrada deverá ser atrasado até surgir o próximo TS2, o que só acontece na trama seguinte. Admitindo que a via de entrada do TSI é um E1, o atraso introduzido seria de (32-7)+2=27 *time slots*.

Como a comutação digital é feita a 4 fios é necessário separar dentro do comutador o percurso de ida do percurso de vinda. Usando a estrutura de *time-slots* apresentada na Figura 6.18, exemplifica-se na Figura 6.19 um caso de comutação temporal bidireccional. Uma chamada é comutada entre os TS2 (de ida e volta) da via PCM/TDM *A* e os TS7 (de ida e volta) da via PCM/TDM *B.* Como já se viu anteriormente, a conexão de A_i/TS2 (da via de ida de A) para B_v/TS7 (da via de vinda de B) requer um atraso de 5 *time-slots,* enquanto a conexão de B_i/TS7 para A_v/Ts2 requer um atraso de 27 *time-slots.* No estudo seguinte vai-se considerar unicamente uma direcção de transmissão.

O elemento básico de um TSI é uma **memória de dados** onde as palavras PCM correspondentes aos diferentes *time-slots* de uma trama são armazenadas (atrasados) temporariamente. O intervalo de tempo em que essas palavras são guardadas na memória depende do atraso a introduzir e poderá ir de um *time-slot* até uma trama menos um *time-slot.*

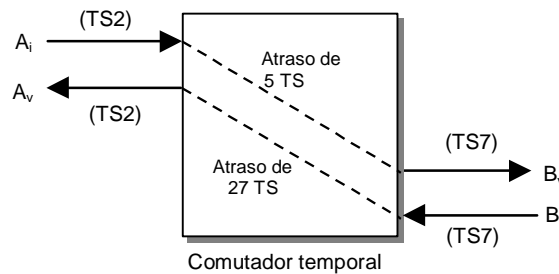



Figura 6.19 Comutação temporal bidireccional.

A escrita das palavras PCM na memória de dados pode ser controlada por um simples contador. Cada vez que o contador é incrementado, uma nova palavra PCM é armazenada numa posição sequencial na memória de dados (escrita sequencial). Deste modo, o conteúdo do *time-slot* 1 é escrito na célula 1, o conteúdo do time-slot 2 na célula 2, e assim sucessivamente. A leitura da memória de dados é controlada pela **memória de endereços**. Esta memória deverá ter um número de células igual ao número de *time-slots* por trama e a dimensão de cada célula deverá ser suficiente para endereçar todas as células da memória de dados. Admitindo que a trama da via TDM (de entrada e saída) possui *w time-slots*, cada um com 8 bits, tem-se que a dimensão da memória de dados deverá ser $8\times w$ bit, enquanto a dimensão da memória de endereços deverá ser $w\times\log_2 w$ bit.

A Figura 6.20 mostra um sistema de comutação temporal simples fazendo uso de *5 time-slots*. A escrita na memória de dados é controlada por um contador. O conteúdo do

TS1, a palavra A, é escrito na célula 1. De modo idêntico, o conteúdo do TS2, a palavra B, é escrita na célula 2, e assim, sucessivamente. Os endereços de leitura são armazenados na memória de endereços de acordo com a ordem requerida para a sequência de saída. No caso em análise, o conteúdo da célula 5 (a palavra E) da memória de dados deverá ser lida durante o time-slot 1, o que explica que a primeira célula da memória de endereços tenha armazenado o endereço "5". A segunda célula da memória de endereços contém o endereço "4", indicando que a célula 4 da memória de dados deverá ser lida durante o *time-slot* 2 de saída. A mesma lógica de raciocínio aplica-se às outras células.

Em operação, em cada *time-slot* 1 o conteúdo do TS1 da via TDM de entrada (palavra A) é escrito na primeira célula de memória de dados, usando o endereço gerado pelo contador. No mesmo *time-slot* o conteúdo da célula 5 da memória de dados (palavra E) é lida usando o endereço armazenado na memória de endereços. Durante os *time-slots* 2, 3, 4, e 5 são lidas as palavras PCM D, A, C e B. Assim, o conteúdo do TS1 da via TDM de entrada sofre um atraso de 3-1=2 *time-slots*. De modo similar, poder-se-ia calcular os atrasos sofridos nos outros *time-slots*. Um aspecto relevante a reter é que, durante cada *time-slot*, tem lugar uma operação de escrita e leitura na memória de dados. Para evitar problemas de conflitos de escrita/leitura nos casos em que o atraso é nulo, a operação de escrita ocorre na primeira metade do *time-slot,* e a operação de leitura é executada na segunda metade do *time-slot*, como se mostra na Figura 6.20.

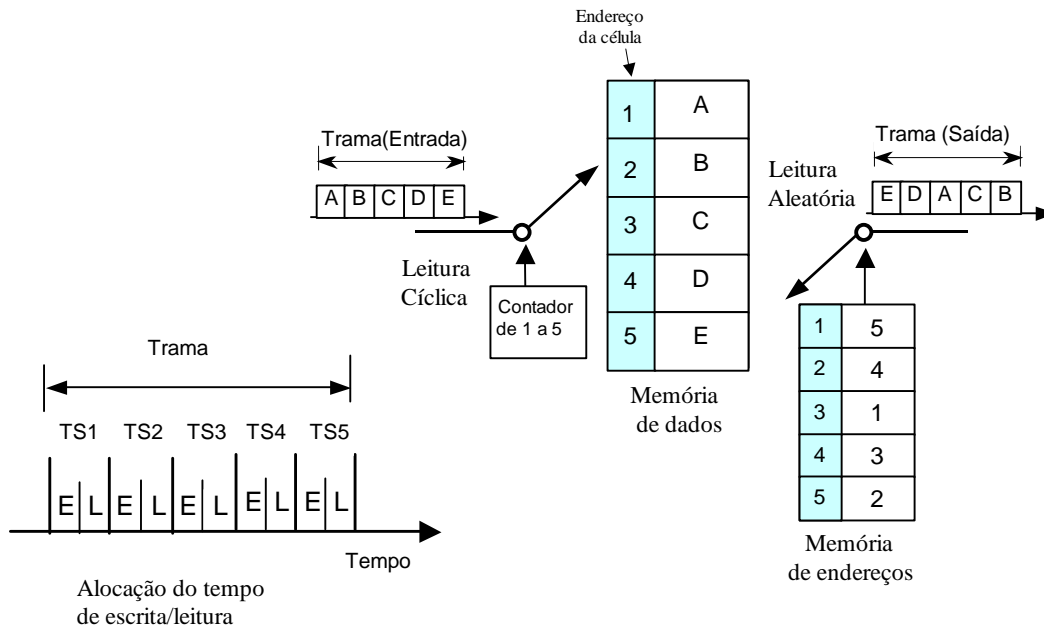



Figura 6.20 Princípio de funcionamento de um TSI.

A sequência de procedimentos descrita repete-se trama a trama, até que o conteúdo da memória de endereços seja alterado pelo sistema de controlo da central de comutação, o que acontece sempre que haja uma terminação ou o estabelecimento de uma nova chamada.

Saliente-se, que a dimensão máxima de um sistema de comutação temporal é limitada pelo tempo de acesso à memória. Seja w o número de canais (número de *time-slots*) da via TDM de entrada, D_b o débito binário por canal e b o número de bits em cada time-slot. A duração máxima de cada *time-slot* é $b/(D_b w)$. Atendendo ao facto que durante um *time-slot* tem lugar uma operação de leitura e uma operação de escrita, o tempo de acesso à memória é dado por

$$t_a \leq \frac{b}{2wD_b} \tag{6.56}$$

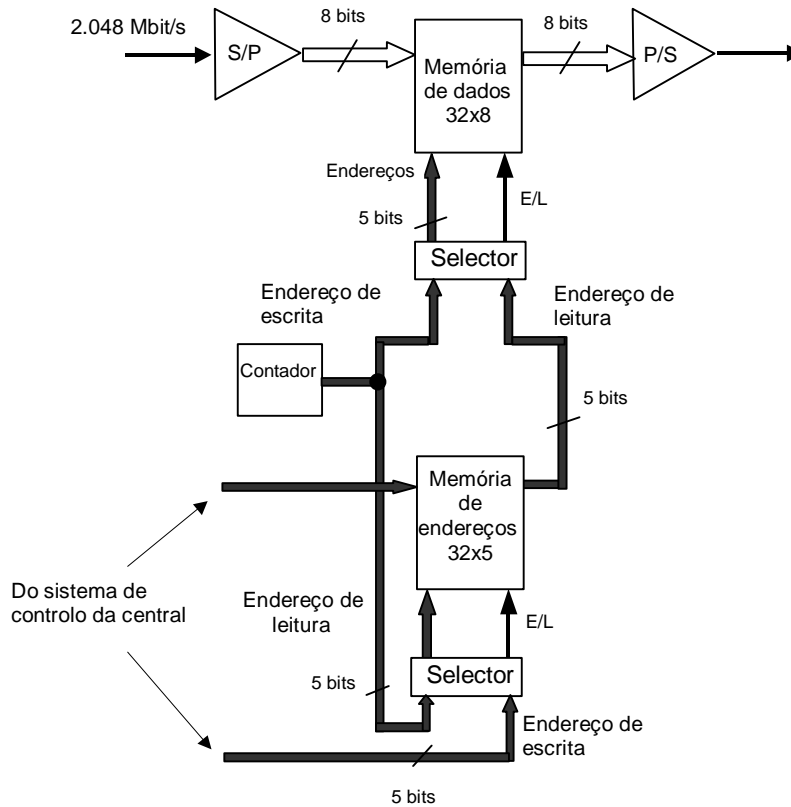



Figura 6.21 Estrutura básica de um TSI de 32 canais.

Exemplo 6.6
Para uma via TDM de entrada de um TSI com 32 canais (64 kb/s por canal de 8 bits) o tempo de acesso deverá ser $t_a \leq 8/(2\times32\times64\text{x}10^3)$=1.95 μs. Para uma via com 128 canais tem de ser $t_a \leq$ 488 ns. Por sua vez, uma via TDM com 16 384 canais requer uma memória com um tempo de acesso igual ou inferior a 3.8 ns, o que é impraticável.

Depois de se analisarem os princípios básicos de operação de um TSI vale a pena tecerem-se algumas considerações sobre aspectos práticos de implementação desses comutadores. A Figura 6.21 apresenta a estrutura básica de um TSI de 32 canais. Nesta estrutura, quer a memória de dados, quer a memória de endereços são implementadas usando RAMs (*random-access-memory*), operando em modo paralelo. Cada RAM tem

uma entrada, uma saída, um porto de endereços, um porto de comando de escrita/ leitura (E/L) e uma entrada de relógio (não representada na Figura 6.21). O sinal de relógio é extraído do sinal E1 (2.048 Mbit/s) que, como se sabe. transporta 32 canais PCM (8 bits), dos quais 30 são canais de informação propriamente dita. Como consequência, ambas as memórias têm 32 células requerendo palavras de endereçamento com 5 bits.

Como já se referiu, anteriormente, durante cada *time-slot* as palavras PCM são primeiro escritas e depois lidas da memória de dados. O selector é usado para controlar se são os endereços de escrita ou leitura que são aplicados no porto de endereços e ainda para gerar o comando de escrita/leitura. O funcionamento do TSI é idêntico ao já descrito. A sequência binária passa por um conversor série/paralelo (S/P), que forma palavras de 8 bits. Estas palavras são escritas sequencialmente na memória de dados (dimensão 32×8) sobre o controlo de um contador e são lidas aleatoriamente sobre o controlo da memória de endereços. O sinal na saída passa por um conversor paralelo/série para gerar a sequência binária de saída. O padrão de conexões a ser implementado pelo TSI é estabelecido pelo sistema de controlo da central, escrevendo na memória de endereços os endereços apropriados. Esta operação só ocorre quando há chamadas para estabelecer ou terminar. Em todas as outras situações a memória de endereços permanece no estado de leitura.

6.2.5 Arquitecturas de comutação digital

6.2.5.1 Arquitecturas TS

As centrais digitais usam, normalmente, arquitecturas de comutação baseadas na combinação de matrizes espaciais e TSIs. Na Figura 6.22 representa-se uma estrutura constituída por dois andares: um andar temporal (andar *T*) constituído por vários blocos TSI, seguido de um andar espacial (andar *S*) consistindo numa matriz espacial. Esta estrutura designa-se por comutador tempo-espaço, ou TS.

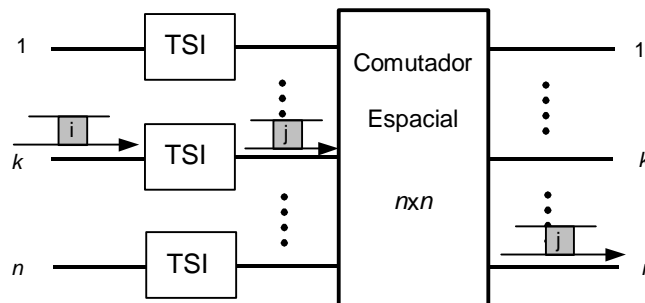



Figura 6.22 Princípio de funcionamento de um comutador TS.

Para se analisar o funcionamento desta estrutura pode-se considerar que cada canal é caracterizado pelo par ordenado dos índices de linha e *time-slot.* Assim, o par (k,i) identifica o canal correspondente ao *time-slot* i da linha k. A acção do comutador, no caso exemplificado, pode ser vista como uma troca do par (k,i) pelo par (n,j). O TSI é

responsável pela troca do índice do *time-slot* $(i \rightarrow j)$. Por sua vez, durante o *time-slot j,* a matriz espacial assegura o contacto entre a linha de entrada *k* e a linha de saída *n*, ou seja, assegura a troca dos índices $(k \rightarrow n)$.

No tipo de arquitectura em análise ocorre bloqueio sempre que dois utilizadores ocupando diferentes *time-slots* da mesma linha de entrada, pretendem ligar-se ao mesmo *time-slot* em diferentes linhas de saída. A Figura 6.23 mostra um caso em que isso acontece. Admitindo que se encontra activa a ligação entre os canais (1,2) e (*N*,15), a ligação (1,8)→(2,15) é bloqueada já que existe um conflito pela utilização do *time-slot* 15 da linha de saída do primeiro bloco TSI.

O problema do bloqueio referido pode ser ultrapassado adicionando um outro andar de comutação espacial à estrutura TS, de modo a que os intercâmbios dos TS2→TS15 e TS8→TS15 sejam efectuados por TSI diferentes. Esta estrutura é conhecida como um comutador espaço-tempo-espaço (STS). Em alternativa, pode-se considerar uma estrutura tempo-espaço-tempo (TST).

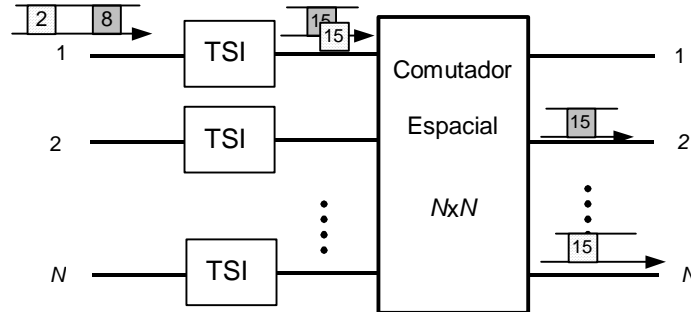



Figura 6.23 Origem do bloqueio num comutador TS.

6.2.5.2 Arquitectura STS

A Figura 6.24 apresenta a estrutura de um comutador STS. Nesta arquitectura existem dois andares de comutação espacial, separados por um andar de comutação temporal.

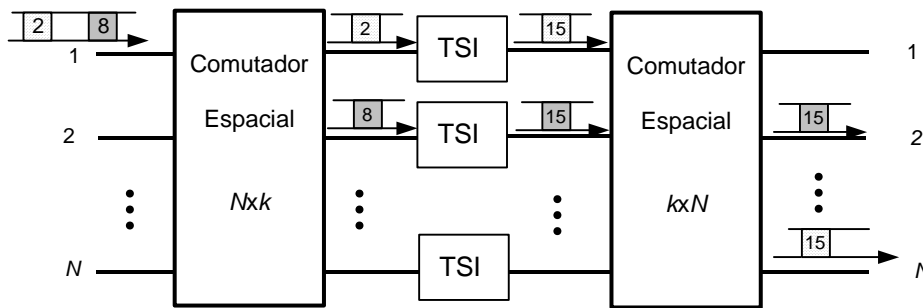



Figura 6.24 Estrutura de um comutador STS.

Como se pode verificar o conflito associado à utilização do *time-slot* 15 foi resolvido. O comutador espacial de entrada encarrega-se de ligar no TS2 a entrada 1 à saída 1 e no TS8 a entrada 1 à saída 2. O primeiro bloco TSI transfere o conteúdo do TS2 para o TS15, enquanto o segundo bloco transfere o conteúdo do TS8 para o TS15. O andar espacial terminal é responsável por estabelecer no TS15 as ligações apropriadas.

Normalmente, os comutadores espaciais são implementados usando estruturas monoandar (sem bloqueio), embora para comutadores de grandes dimensões se possam considerar comutadores espaciais com andares múltiplos. Quando cada andar individual é uma estrutura sem bloqueio, a operação do comutador STS é funcionalmente idêntica à operação de uma matriz espacial com três andares. Assim, para estudar a probabilidade de bloqueio, pode-se usar um grafo de canal idêntico ao da Figura 6.15 (ver Figura 6.25). Atendendo a esse grafo a probabilidade de bloqueio vem dada por

$$B = \left(1-(1-p_1)^2\right)^k \qquad (6.57)$$

onde k é o número de blocos TSI no andar intermédio e p_1 é a probabilidade de ocupação de cada linha intermédia (entre os andares espaciais e o andar temporal). Esta probabilidade é facilmente calculada em função da operação de expansão ou compressão realizada no primeiro andar do comutador, vindo

$$p_1 = p\frac{N}{k} \qquad (6.58)$$

onde N representa o número de entradas do andar espacial de entrada. Para desenhar um comutador STS sem bloqueio em sentido estrito deverá aplicar-se o teorema de Clos, definido anteriormente para as estruturas espaciais com três andares, o que conduz a [Be91]

$$k \geq 2N - 1 \qquad (6.59)$$

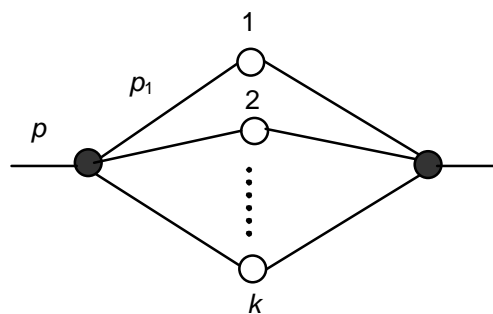



Figura 6.25 Grafo de canal de um comutador STS.

6.2.5.3 Arquitectura TST

A estrutura de um TST é dual da anterior, consistindo em dois andares temporais separados por um andar espacial (ver Figura 6.26). A informação correspondente a um canal de entrada é atrasada no bloco TSI de entrada, até encontrar um *time-slot* livre no andar espacial. Durante esse *time-slot* a informação é transferida através da matriz espacial para o TSI de saída, onde fica armazenada até ser enviada para a saída do comutador no *time-slot* apropriado.

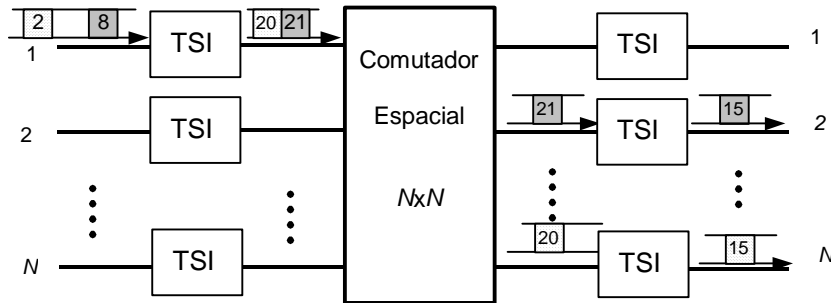



Figura 6.26 Estrutura de um comutador TST.

Como já se referiu anteriormente, o padrão de interligações de uma matriz espacial digital altera-se de *time-slot* para *time-slot.* Além disso, o número de time-slots da matriz espacial digital (*time-slots* internos), *l*, é independente do número de *time-slots* das vias TDM de entrada e saída (*time-slots* externos). Atendendo a esses factores, uma matriz espacial digital pode ser vista, sobre o aspecto funcional, como um conjunto de *l* camadas paralelas, onde cada camada corresponde ao padrão de interligações associado a um determinado *time-slot* [Mc90]. Na Figura 6.27, mostra-se, por exemplo, para uma matriz 4×4 o padrão de interligações no *time-slot* 1 (camada 1) e no time-slot *l* (camada *l*). A partir destas considerações é fácil concluir que para garantir o mesmo grau de conectividade requerer-se-iam *l* matrizes analógicas com a mesma dimensão da matriz digital.

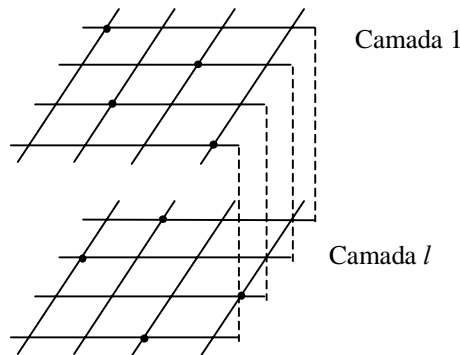



Figura 6.27 Equivalente analógico de uma matriz espacial digital.

Este comportamento da matriz espacial digital faz com que o grafo de canal do comutador TST, seja idêntico ao das matrizes espaciais com três andares, com l sub-matrizes no andar intermédio (ver Figura 6.28).

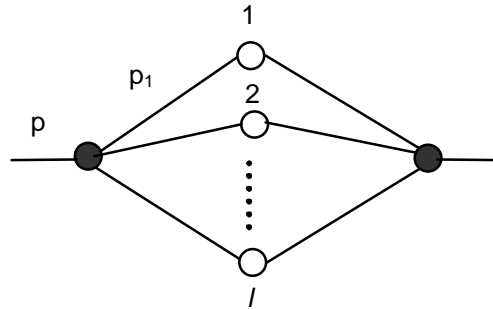



Figura 6.28 Grafo de canal de um comutador TST.

Designando por w o número de *time-slots* (canais) externos, e tendo presente que nenhum dos andares individuais do comutador TST introduz bloqueio, tem-se a seguinte expressão para a probabilidade de bloqueio de interligação da estrutura em análise [6,7]:

$$B = \left(1 - (1 - p_1)^2\right)^l \qquad (6.60)$$

onde

$$p_1 = p\frac{w}{l} \qquad (6.61)$$

O bloqueio de interligação no comutador TST ocorre somente quando não há *time-slots* internos disponíveis para estabelecer uma ligação entre o andar temporal de entrada e o andar temporal de saída. Certamente que, a probabilidade de bloqueio ir-se-á reduzir se o número de time-slots internos l aumentar, e por analogia com o caso analógico, tem-se que para que o comutador TST não introduza bloqueio dever-se-á verificar a condição

$$l \geq 2w - 1 \qquad (6.62)$$

6.2.6 Considerações finais

Em síntese, pode-se referir que, para comutadores de pequenas dimensões (até 250 circuitos) pode-se usar um único andar T (um único andar S não é apropriado devido a introduzir um bloqueio elevado). Os comutadores de capacidade mais elevada, como aqueles que são requeridos pelas centrais de comutação telefónica, resultam da combinação de vários andares S e T. As estruturas com dois andares, ST ou TS, permitem realizar comutadores com capacidades médias, mas introduzem bloqueio, cuja probabilidade aumenta com o aumento das suas dimensões. Por conseguinte, estruturas com médias ou altas capacidades e probabilidades de bloqueio reduzidas requerem no mínimo três andares.

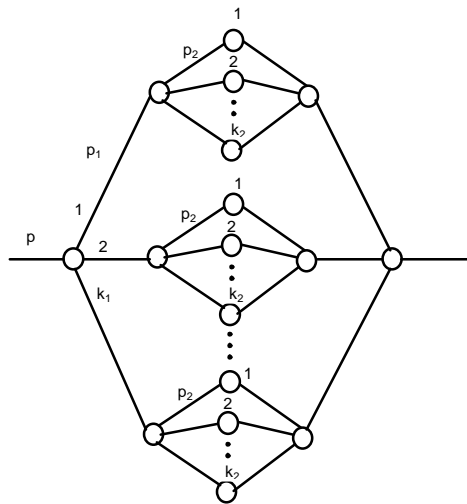
As primeiras centrais digitais usavam uma arquitectura STS. Contudo, a partir do fim da década de 70, a arquitectura TST tornou-se predominante e hoje é usada invariavelmente nos comutadores digitais multiandar. A utilização inicial das estruturas STS devia-se ao custo elevado das memórias e à necessidade de minimizar a componente temporal. Com o desenvolvimento das memórias RAM de alta-velocidade o panorama alterou-se, sendo hoje em dia, mais barata a comutação temporal do que a espacial.

As dimensões de um comutador espacial digital aumentam com o quadrado do número de linhas de entrada/saída. Deste modo, para diminuir as suas dimensões, usa-se a mesma filosofia já descrita para o caso analógico, ou seja, as matrizes são divididas em vários andares, o que conduz a estruturas do tipo TSST ou TSSST.

6.2.7 Problemas

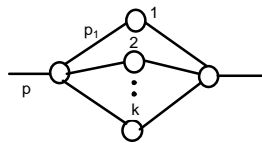
6.12 Considere um sistema de comutação espacial com 3 andares, com 2048 entradas 2048 saídas, que não apresenta bloqueio. Numa configuração optimizada, qual é o número de pontos de cruzamento necessários? Qual o número de matrizes em cada andar e quantas linhas de entrada e saída existem por andar?

6.13 a) Determine, recorrendo ao método dos grafos de Lee, uma expressão para a probabilidade de bloqueio de uma arquitectura de comutação com três andares.

b) Determine uma expressão para a probabilidade de bloqueio de uma arquitectura de comutação com 5 andares, como se indica na figura, recorrendo à mesma metodologia.

6.14 Um comutador TSI possui uma memória com um ciclo de acesso de 50 ns (correspondente a uma operação de escrita ou de leitura). Cada palavra de memória tem 16 bits e o comprimento das tramas é de 1024 bits. Qual é débito binário máximo de cada canal? Qual é o débito binário nas vias de entrada e de saída? Quantos bits de memória são necessários nesta unidade? Qual é o atraso máximo introduzido pelo TSI?

6.15 Um comutador espacial digital é projectado para comutar canais PCM de 64 kbit/s. Os canais que são codificados usando 8 bits/amostra e agregados em sinais multiplexer com 32 *time-slots.* O comutador dever ter capacidade para 16 deste sinais (512) canais. Desenhe o comutador, apresentado os requesitos para todas as memórias e frequência de relógio.

6.16 Considere uma arquitectura de comutação TST com 10 linhas de entrada e 100 canais por linha. Sabendo que a probabilidade de ocupação dos canais de entrada é de 0.1 e que a estrutura possui 20 *time-slots* internos, calcule a probabilidade de bloqueio de interligação.

6.17 Considere um sistema de comutação para 2500 utilizadores descrito pelo seguinte grafo de Lee (k=10, p=5%, 8 bits/*time-slot* e concentração do andar de entrada de 10):

a) Calcule a probabilidade de bloqueio do sistema.
b) Sabendo que o sistema é do tipo STS, desenhe a sua estrutura e dimensione cada um dos andares de comutação.

6.2.8 Referências

[Be62] V. E. Benes, "On rearrangeable three-stage switching networks," *Bell Syst*. *Tech. Jour*., 41, pág. 1481-92, 1962.

[Be91] J. Bellamy, *Digital Telephony*, John Wiley & Sons, Inc., 1991.

[Cl53] C. Clos, "A study of non-blocking switching networks," *Bell System Technical Journal*, Março 1953, pág. 406-424.

[Fl95] J. E. Flood, *Telecommunications Switching, Traffic and Networks*, Prentice Hall, 1995.

[Le55] Lee, C. Y., "Analysis of switching networks," *Bell System Techn. Journal*, 34, pág.1287-1315, 1955.

[Mc90] J. C. McDonald, *Fundamentals of Digital Switching*, Plenum Press, 1990.

[Re95] F. J Redmill e A.R.Valdas, *SPC Digital Telephone Exchanges*, Peter Peregrinus, 1995.

6.3 Sinalização

6.3.1 Objectivos e tipos de sinalização

A sinalização pode ser vista como o sistema nervoso de uma rede de telecomunicações, transformando redes inertes, sem vida, em máquinas dinâmicas poderosas capazes de proporcionarem diferentes serviços aos utilizadores.

A sinalização é responsável pela transferência de informação de controlo entre os utilizadores e a rede (**sinalização de assinante**), e entre as centrais da rede (**sinalização de troncas**), sendo, assim, responsável por estabelecer, manter e remover qualquer ligação entre dois assinantes, sendo também uma peça fundamental na gestão da rede.

Um dos factores que mais tem influenciado o desenvolvimento dos sistemas de sinalização é a relação entre as funções de sinalização e as de controlo nas centrais de comutação. No caso das centrais analógicas as funções de controlo, que eram realizadas usando relés telefónicos, estavam intimamente ligadas às funções de comutação. Neste caso, os caminhos físicos de sinalização e de informação (voz) são os mesmos, daí designarem-se por sistemas de **sinalização de canal associado** ou CAS (*channel-associated signalling*).

A fase seguinte de evolução consistiu em separar as funções de controlo das funções de comutação. Esta técnica permite usar computadores para realizar as funções de controlo, permitindo maior flexibilidade e redução nos custos. As centrais que usam controlo por computador designam-se por **Centrais SPC** (*stored program control*). Neste caso, a sinalização e o tráfego de voz continuam a partilhar o mesmo caminho, embora sejam separados dentro da central, continuando, deste modo a ser sistemas de sinalização de canal associado. Isto é ilustrado na Figura 6.29, onde o circuito de voz (a cheio) é encaminhado pela matriz de comutação, mas a sinalização é encaminhada pelo sistema de controlo. Entre as centrais A e B a sinalização e o tráfego seguem o mesmo caminho.

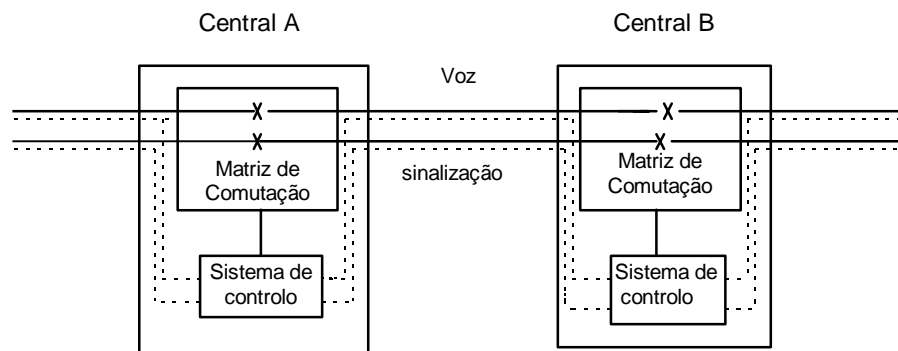



Figura 6.29 Sistema de sinalização de canal associado com blocos de controlo e sinalização separados.

Um segundo tipo de sistema de sinalização é denominado de **sinalização em canal-comum** ou CCS (*common-channel signalling*). Num sistema deste tipo é proporcionado um caminho comum para um determinado número de circuitos de sinalização, o que tem subjacente a separação dos caminhos de voz e sinalização. Esta

separação ocorre, quer no interior, quer no exterior da central, como se exemplifica na Figura 6.30. Essa figura mostra que no ambiente CCS os circutos de voz são encaminhados pela matriz de comutação como anteriormente. Contudo, a sinalização (a tracejado) é encaminhada por um caminho diferente, tanto no interior como no exterior da central. Esta solução é particularmente vantajosa quando adoptada em paralelo com comutação e transmissão digital, como é o caso das redes digitais integradas. O aproveitamento completo dos benefícios dos sistemas CCS só terá, contudo, lugar nas redes digitais com integração de serviços (RDIS).

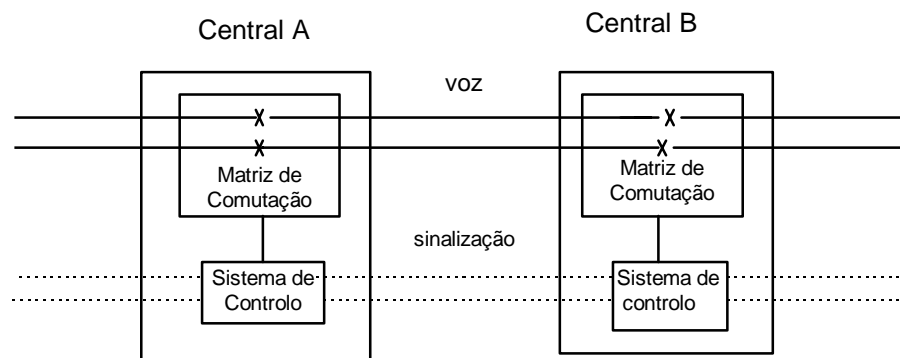



Figura 6.30 Sinalização em canal comum.

6.3.2 Funções da sinalização

Numa rede telefónica moderna as funções gerais da sinalização são as seguintes:

- **Alerta:** Tem a ver com os pedidos de serviço e com o progresso da ligação. O pedido inicial de um assinante, manifestado no levantamento do auscultador é uma função de alerta. Também a central local envia sinais de alerta para outras centrais para requerer linhas de troncas. A corrente de toque da campainha, enviada pela central local para o telefone destinatário, desempenha, também, uma função de alerta. Esta corrente tem uma frequência de 25 Hz e está 2s em estado ligado e 4s em estado desligado.

- **Endereçamento**: O endereço, ou seja, o número telefónico do destinatário deve ser transmitido pelo assinante chamador. Esta função pode ser realizada usando, quer pulsos de marcação, quer conjuntos de duas frequências puras (tons).

- **Supervisão**: As centrais de comutação necessitam de saber se as linhas estão inactivas ou em uso. Estas funções podem ser realizadas pela presença ou ausência de corrente na linha, ou então, pela presença ou ausência de uma determinada frequência (note-se o seu carácter binário). Também é necessário saber quando um determinado circuito pode ser libertado, o que também constitui uma função da supervisão.

- **Informação**: O sinal de linha, o sinal de linha ocupada, o sinal de toque da campainha transmitido em sentido inverso, os diferentes anúncios gravados enviados para o assinante são sinais de informação.

- **Tarifação:** Informação usada para efeitos de facturação.

- **Gestão da rede:** Sinais usados para efeitos de manutenção, diagnóstico e operação.

O sinal de linha é uma combinação das frequências de 350 e 440 Hz. O sinal de linha impedida corresponde à transmissão das frequências de 480 Hz e 620 Hz, estando ligado 0.5 s e desligado o mesmo tempo. O sinal de tronca ocupada (enviado quando a central ou a linha de troncas estão ocupadas) corresponde às mesmas frequências que o caso anterior, mas com uma frequência de repetição dupla. O estado ligado e desligado duram 0.25s.

6.3.3 Sinalização de assinante

A sinalização de assinante deve ser o mais simples possível, sendo tradicionalmente obtida pela presença ou ausência de corrente no lacete de assinante.

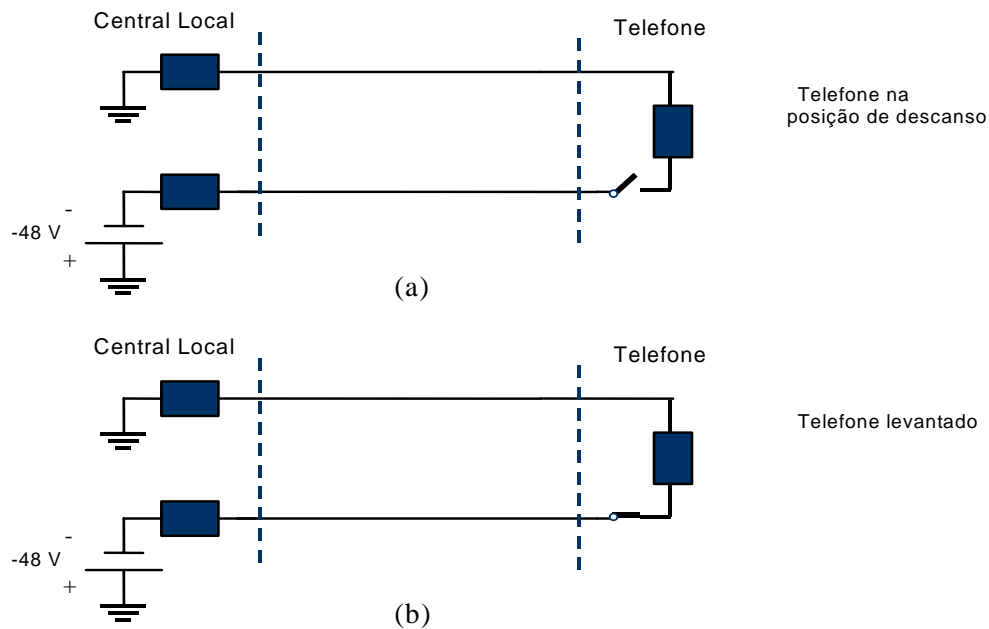



Figura 6.31 Lacete do assinante. (a) Corrente da linha reduzida; (b) Corrente elevada.

O telefone é alimentado por corrente contínua a partir da central local, com baterias que geram uma tensão de –48 V. Quando o auscultador está na posição de descanso, a corrente que circula na linha do assinante é nula (ver Figura 6.31). Quando o auscultador é levantado fecha-se o lacete de assinante e por conseguinte estabelece-se um fluxo de

corrente na linha de assinante. O início da circulação da corrente é detectado pelo ILA da central local (relé da linha) a qual identifica a linha que a originou. A campainha está sempre ligada à linha telefónica. A capacidade em série com a campainha (ver Figura 4.10) evita que esta seja percorrida pela corrente contínua.

A frequência da corrente de toque da campainha correspondente ao **sinal de chamada** é de cerca de 25 Hz e a sua tensão *rms* é de 75 V. Essa corrente está ligada 2 e deligada 4s. Quando o assinante chamado responde, a sua central local detecta a circulação da corrente contínua e desliga a corrente de toque da campainha.

6.3.4 Tipos de sinalização de endereçamento

Para efectuar uma ligação telefónica é necessário enviar para a central de comutação local o endereço (número) do destinatário. Esta informação, pode ser enviada por meio de pulsos decádicos correspondentes à interrupção da corrente contínua, ou por meio de sinalizacão multifrequência. A sinalização decádica é usada nos **telefones de disco** e a sinalização **multifrequência de dois tons** nos **telefones de teclado.**

Sinalização decádica

O disco do telefone possui dez furos, correspondentes aos 10 algarismos que são usados para a transmissão da informação. Inicialmente, o disco é girado no sentido dos ponteiros do relógio. Durante o retorno à sua posição de repouso as molas de pulso vão interrompendo a ligação, tantas vezes quanto o valor do algarismo marcado. No caso da Figura 6.32, verifica-se que existem quatro interrupções da corrente na linha, correspondendo portanto ao número 4.

Figura 6.32 Sinalização decádica.

Notar que, o atraso de marcação depende dos dígitos marcados. Um zero requer 10 pulsos e como cada pulso requer 100 ms, necessita-se de 1 s para marcar um zero.

Outro inconveniente da sinalização decádica resulta do facto desses pulsos se distorcerem à mediada que se propagam na linha. Essa distorção é induzida pela capacidade da linha e aumenta com o comprimento da linha. Como o equipamento tem unicamente possibilidade de suportar um determinado grau de distorção será de concluir que este tipo de sinalização contribui para limitar o comprimento do lacete de assinante.

Sinalização multifrequência de dois tons (rede local)

Com a sinalização multifrequência de dois tons todos os dígitos são transmitidos com o mesmo atraso, cerca de 120 ms. A marcação de um número activa dois geradores de

frequências diferentes. Por exemplo, quando o número 2 é marcado os tons de 697 Hz e 1336 Hz são sobrepostos e transmitidos. Para além de requerer um tempo de atraso menor e idêntico para todos os números, esta técnica usa tecnologia do estado sólido, a qual é compatível com os comutadores electrónicos. Este tipo de sinalização requer, contudo, o uso de um receptor apropriado do lado da central, aumentando por conseguinte o custo do sistema.

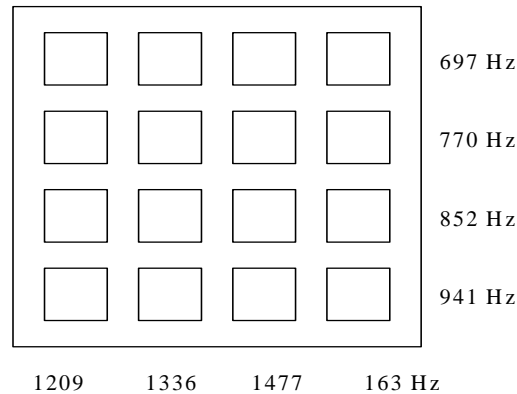



Figura 6.33 Frequências usadas na sinalização multifrequência.

6.3.5 Sinalização na rede de troncas

As distâncias de transmissão em jogo na rede de troncas (rede entre as centrais de comutação) são normalmente elevadas. Assim, no caso de se usar a sinalização DC (presença ou ausência de corrente) os pulsos para além de serem distorcidos, iriam ser consideravelmente atenuados. Para compensar essa atenuação requer-se-ia a utilização de amplificadores DC, que têm a desvantagem de serem muito caros. Além disso, requer-se um meio de transmissão metálico, tipo linha telefónica, o qual não pode ser usada para a transmissão de sinais multiplexados. Essas limitações impedem a utilização da sinalização DC na rede de troncas.

Como alternativa usa-se a sinalização AC, que é baseda na transmissão de sinais multifrequência (MF). A sinalização AC pode ser ***na banda***, sempre que as frequências de sinalização sejam transmitidas na banda do canal telefónico (300-3400 Hz), ou ***fora da banda*** no caso em que essa transmissão ocorre entre os 3400-4000Hz. A maioria dos sistemas de sinalização usa o primeiro tipo. Os sistemas fora da banda apresentam a vantagem de a sinalização e a voz poderem ser transmitidos simultaneamente, mas requerem uma maior banda, que nem sempre está disponível. Daí só se usarem nos sistemas mais antigos baseados em FDM.

Como exemplo de sinalização na banda refira-se um esquema usado vulgarmente, que transmite uma frequência de 2600 Hz para indicar que a tronca está inactiva. A partir do momento que a tronca passa para um estado activo esse sinal é desligado. Quanto à informação de endereçamento, esta é enviada usando uma sinalização multifrequência parecida com a usada nos telefones de teclas. Os tons usados são os de 700, 900, 1100,

1300, 1500 e 1700 Hz. Qualquer dígito é transmitido usando dois desses tons em diferentes combinações.

Note-se que a sinalização multifrequência é independente do meio de transmissão, podendo usar-se com cabos de pares simétricos, cabos coaxiais, fibra óptica, etc. Porém, a sinalização por interrupção da corrente DC pode ser unicamente usada com cabos de pares simétricos, e com um fonte de alimentação contínua de -48V, como é o caso do lacete de assinante.

Outro tipo de sinalização de troncas é a sinalização PCM, que já foi analisada quando se estudou a multiplexagem TDM. Como se viu, na primeira hierarquia do ITU-T usa-se o *time-slot* 16 para transmitir a informação de sinalização. Em cada trama transmite-se a informação de sinalização de dois canais. Para transmitir a informação dos 30 canais são necessárias 15 tramas. Adicionando mais uma para efeitos de enquadramento, tem-se uma multitrama constituída por 16 tramas.

Os tipos de sinalização referidos pertencem todos à família dos sistemas de **sinalização em canal associado,** já que se usa o mesmo meio físico para transmitir o sinal de voz e o de sinalização. Como já se viu, este tipo de sinalização ainda se pode dividir em sinalização **na banda** e **fora da banda**. Notar que, esta última usa o mesmo meio físico que a voz, mas uma diferente banda. Pode, assim, ser vista como uma multiplexagem de frequência dentro do circuito de voz. A sinalização usando a presença ou ausência de uma corrente DC pode ser integrada neste tipo de sinalização.

Como já se viu, em alternativa tem-se **a sinalização em canal comum**. Neste caso, tem-se uma rede distinta para transmitir os sinais de voz e os sinais de sinalização. Para além disso, a informação de sinalização é digitalizada, podendo-se em alguns casos usar o canal 16 da trama E1 para transmitir essa informação. Só que, nestes sistemas a estruturação em multitrama não é relevante. A capacidade dentro desse canal é atribuída dinamicamente, de acordo com as necessidades dos diferentes canais. Por exemplo, se o canal de voz 30 necessitasse de transferir informação de sinalização, seria usada toda a capacidade do canal 16 até que essa transferência ficasse completa. Depois de completa, essa capacidade seria atribuída a outro canal.

6.3.6 Análise do processmento de uma chamada telefónica

As etapas envolvidas no processamento de uma chamada local são as seguintes:

1) O assinante chamador começa por levantar o microauscultador, o que imediatamente contribui para estabelecer uma corrente DC na linha de assinante (**sinal de apresar** ou *seize* segundo a designação inglesa).

2) A presença da corrente DC é entendida pela interface de linha de assinante da central local como um pedido de serviço, e o assinante que requereu esse pedido é identificado. Essa identificação é necessária para efeitos de tarifação, para conhecer o tipo de terminal (teclas ou disco) e ainda para descobrir se é permitido ao assinante fazer a chamada. Essa informação é armazenada nas centrais SPC numa área de memória designada por **registo de chamada**.

3) Depois de disponibilizar uma área de memória para funcionar como registo dos dígitos é enviado para o assinante chamador o **sinal de linha**, indicando que a central está pronta para receber os dígitos do endereço do telefone destinatário.

4) Depois de receber o sinal de linha o assinante procede à marcação dos dígitos. Os dígitos são enviados como sinais para a central e são aí armazenados.

5) O sistema de controlo da central analisa os dígitos do endereço e escolhe o circuito de saída apropriado para encaminhar a chamada. Se esse circuito está ocupado, a central envia o sinal de linha impedida para o assinante chamador.

6) Neste ponto o sistema de controlo conhece a identidade, quer do circuito de entrada, quer do de saída. A fase seguinte consiste em estabelecer um caminho entre eles através da matriz de comutação da central.

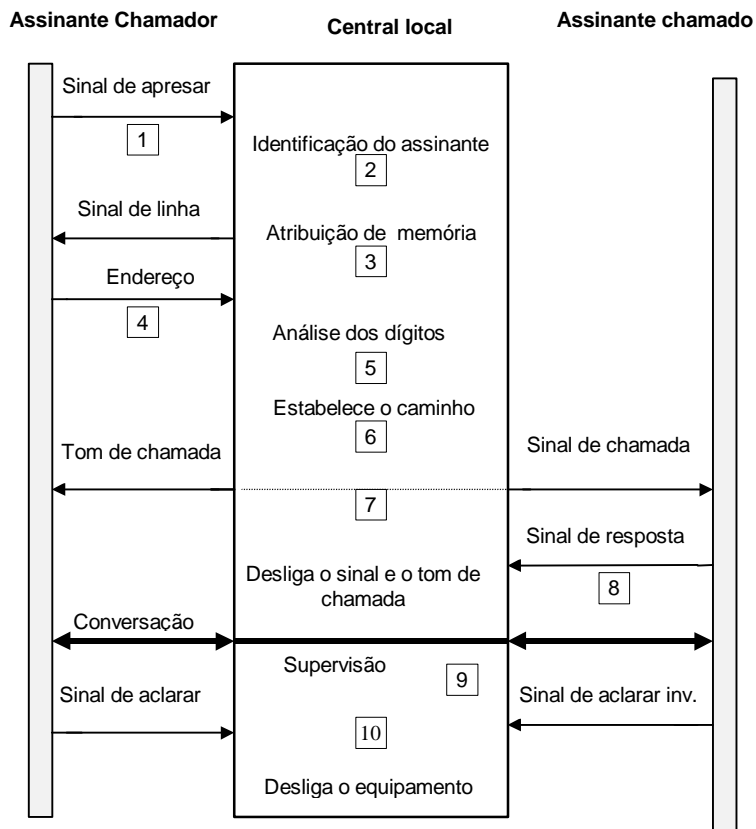



Figura 5.34 Etapas associadas à realização de uma chamada local.

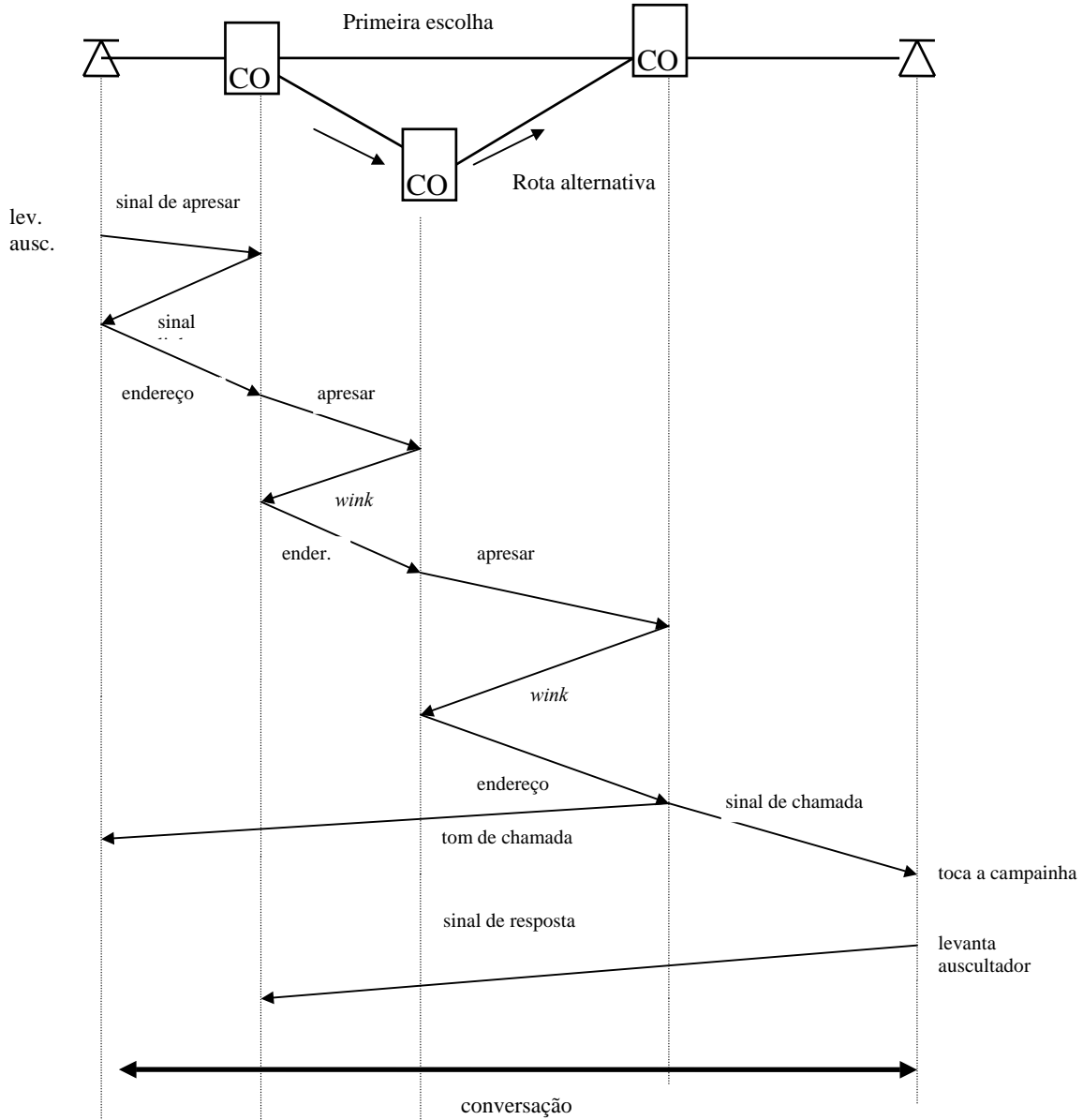



Figura 5.35 Etapas associadas ao processamento de uma chamada que faz uso de três centrais.

7) O **sinal de chamada** é enviado para o destinatário para o avisar da presença da chamada. Este sinal é constituído por uma corrente AC (25Hz) que activa a campainha do telefone. Simultaneamente é enviado o **tom de chamada** para o assinante chamador.

8) O **sinal de resposta** (estabelecimento da corrente DC) é enviado pelo assinante chamado. A transmissão pode agora iniciar-se através do caminho seleccionado na matriz de comutação. O sinal de chamada e o tom de chamada são removidos e a conversação pode iniciar-se através da ligação entretanto estabelecida.
9) Durante o progresso da conversação tem ainda lugar a supervisão, responsável pela tarifação e pela detecção do momento em que um dos assinantes desliga.

10) Quando o **sinal de aclarar** (ausência de corrente DC no lacete do assinante) directo gerado pelo assinante chamador e o sinal de aclarar inverso gerado pelo assinante chamado são recebidos o equipamento usado na central é libertado e disponibilizado para outras ligações.

Na Figura 6.34 descrevem-se os procedimentos associados à realização de uma chamada local. Quando os dois assinantes envolvidos pertencem a centrais diferentes são necessárias as seguintes funções de sinalização inter-central:

11) A central originária da chamada envia um sinal de apresar para uma linha de troncas livre e pede um registo de dígitos à central que se encontra no extremo da tronca.

12) A central terminal depois de disponibilizar o referido registo envia para a central originária um sinal, que segundo a terminologia americana se designa por *wink*. Este sinal é constituído por um sinal de aclarar seguido de um sinal de apresar.

13) Depois de receber o *wink* a central originária envia o endereço, o qual será armazenado no registo de dígitos. No caso em que na ligação em questão estão envolvidas outras centrais o processo descrito anteriormente repete-se, tantas vezes quanto o número de centrais adicionais, como se verifica na Figura 6.35.

6.3.7 Sinalização em canal comum

6.3.7.1 Estrutura e elementos da rede

O sistema de sinalização em canal comum, desenvolvido pelo antigo CCITT e adoptado como recomendação em 1981, é designado por "*Common Channel Signaling System No. 7*", ou simplesmente SS7.

Este sistema de sinalização, projectado usando os conceitos da comutação de pacotes e estruturado em diferentes níveis conforme o modelo OSI, foi desenvolvido para ser usado tanto em ligações nacionais como internacionais.

A rede do SS7 pode ser vista como uma rede de comutação de pacotes, que é usada para transmitir mensagens de sinalização entre os processadores das várias centrais de comutação (baseadas na comutação de circuitos no caso de rede telefónica).

O sistema de sinalização nº7 define três entidades funcionais**: Ponto de sinalização** ou **SP** (*signaling point*), **ponto de transferência de sinalização ou STP** (*signaling transfer point*) e **via de sinalização** (*signaling link*). Os **SPs** são os nós terminais da rede onde os pacotes, designados por mensagens, são criados ou recebidos, enquanto os **STPs** são comutadores de pacotes responsáveis pelo encaminhamento das mensagens de sinalização entre os vários SPs. As vias de sinalização são ligações de dados capazes de suportar um débito binário de 64 kb/s.

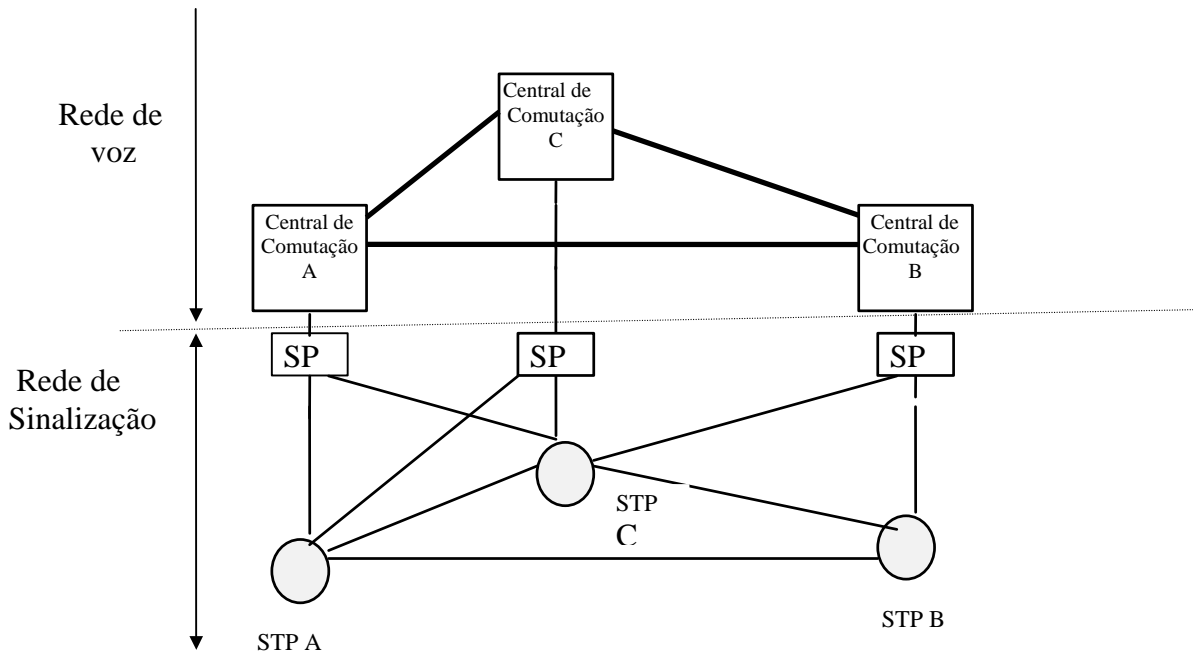



Figura 5.36 Rede de telecomunicações mostrando a rede correspondente ao tráfego de voz e a rede de sinalização.

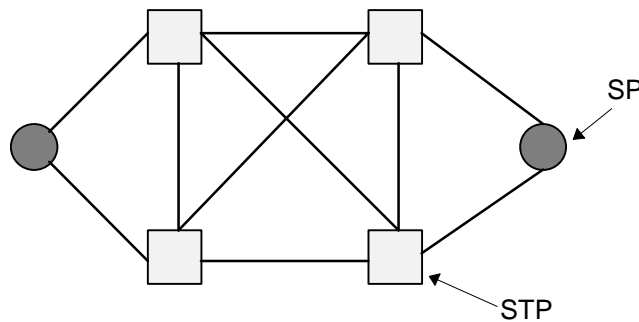



Figura 6.37 Arquitectura básica da rede SS7.

Um das exigências que se coloca a uma rede de sinalização de canal comum é a sua elevada fiabilidade, já que cada via transporta a sinalização de milhares de assinantes (até 30 000). Assim, numa rede SS7 qualquer SP está ligado a dois STP, que por sua vez são implementados aos pares (STP gémeos), e separados geograficamente. A rede é, ainda, dimensionada para usar somente 40% da capacidade de cada STP, de modo a que quando há um falha de um elemento do par o outro STP pode transportar todo o tráfego com uma taxa de ocupação que não ultrapassa os 80%. Os diferentes STP ainda estão ligados entre si com uma topologia em malha, como é ilustrado na Figura 6.37.
Quanto aos SP podem distinguir-se três tipos:

- **Ponto de comutação, ou SP** (*switching point*) o qual é constituido pelo *hardware* e *software* adicionada às centrais de comutação, que é responsável pela conversão do formato das mensagens de sinalização originadas na rede telefónica (Ex: sinalização

DC, sinalização multifrequência, etc.) no formato próprio do SS7. Funciona, assim, como um processador anexo a um comutador existente não SS7, analógio ou digital.

- **Ponto de comutação de serviços ou SSP** (*service switching point*) são centrais de comutação, cujo software é capaz de reconhecer as chamadas que requerem um tratamento especial (acesso a base de dados) antes de serem completadas, como por exemplo, a ligação a serviços inteligentes. Nestas centrais, as funcionalidades associadas ao SS7 residem directamente no interior da própria central digital. As centrais de comutação que possuem unicamente a funcionalidade garantida pelos "*Switching points*", necessitam de recorrer a centrais SSP para acederem a bases de dados.

- **Pontos de Controlo de Serviço ou SCP (***service control points***),** que é o "cérebro" da rede, e que consiste num processador centralizado, que controla a execução dos serviços mais complexos da rede, através do acesso a bases de dados que suportam esses serviços. Ao exercerem o controlo sobre as bases de dados pode dizer-se que estes nós incorporam a inteligência das redes. Tal como os STP, os SCP são implementados aos pares, e cada SCP do par tem armazenado nas suas bases de dados informação idêntica.

Quando um SSP detecta uma chamada especial suspende o procedimento normal e requer a intervenção do SCP. Só depois do SCP enviar a sequência de comandos necessária, é que o referido SSP completa essa chamada. Tem-se assim, que o controlo das chamadas especiais é feito pelo SCP e não pelo SSP.

As bases de dados do SCP têm recursos para processamento da informação, podendo deste modo proporcionarem serviços, tais como, validação dos cartões de crédito, registos para as redes privadas virtuais, tabelas de conversão dos números especiais, etc.

6.3.7.2 Redes inteligentes

A existência dos SCPs na redes SS7 permite a existência de inteligência numa rede, levando necessáriamente ao conceito de **rede inteligente**, cuja configuração genérica se representa na Figura 6.38.

Numa rede com vários SCP deverá existir um sistema de gestão de serviços **SMP** (*service management point*) ligado a todos os SCP através de um "*link*" de dados, que usa o protocolo X25 usado normalmente nas redes de computadores. Este sistema executa a gestão dos nós SCP, gere a adição de novos clientes, actualiza os dados, etc. O facto de haver somente um único SMP, ou em alternativa um número pequeno de SMPs facilita a tarefa da adição de nova informação na rede. Na arquitectura de uma rede inteligente costumam-se definir outros blocos, que contudo não vamos abordar aqui. Para os leitores interessados recomenda-se a referência sobre redes de telecomunicações inteligentes [La95].

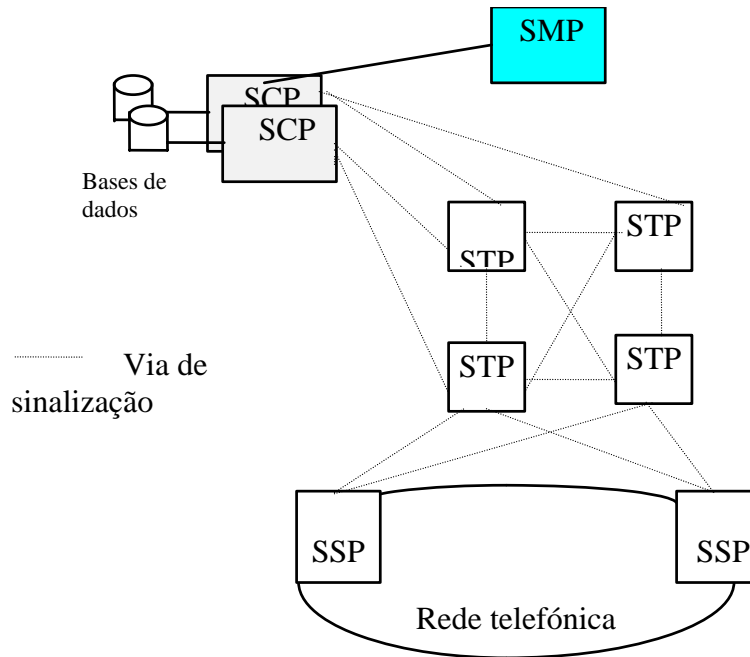



Figura 6.38 Arquitectura genérica de uma rede inteligente.

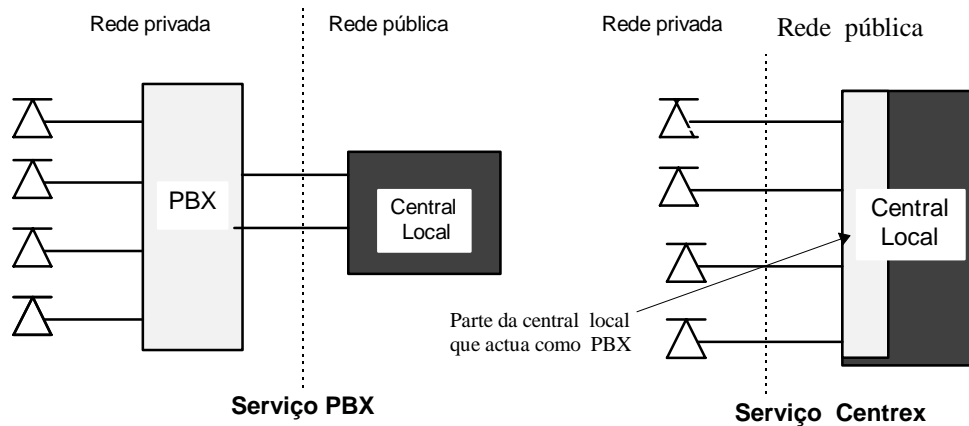



Figura 6.39 Serviço Centrex.

Como exemplos de serviços da rede inteligente refira-se o serviço de chamadas gratuitas (0800 ou 0500), o televoto, a localização de móveis, o serviço Centrex, as redes privadas virtuais etc. A ideia subjacente ao serviço Centrex consiste em proporcionar serviços semelhantes a um PBX, mas usando a central local pública. Esta facilidade oferece aos clientes todas as vantagens dos PBX, sem a necessidade das despesas de investimento e manutenção associadas à sua aquisição. Muitas empresas têm as suas redes próprias usando PBXs. Quando essas redes se estendem por diferentes localidades, em cada localidade têm-se um PBX, os quais estão interligados usando linhas alugadas ao operador de rede pública (usualmente a 2 Mb/s). As redes privadas virtuais proporcionam serviços similares aos dessa rede, mas usando os recursos da rede pública de modo idêntico ao Centrex.

6.3.7.3 Modos de operação

A forma como os diferentes nós da rede de sinalização em canal comum se encontram ligados entre si, permite distinguir três modos de exploração: **associado**, **quase associado** e **não associado**.

No **modo associado** a um dado feixe de circuitos de voz interligando duas centrais de comutação fica afecta uma ligação de sinalização de canal comum, como se exemplifica na Figura 6.40

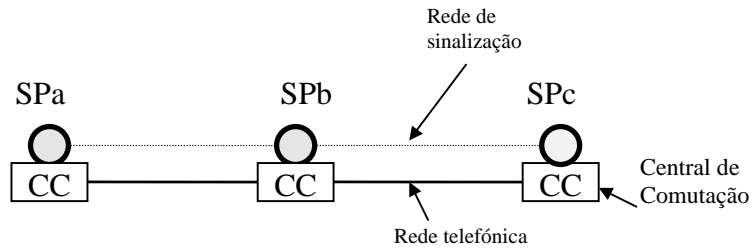



Figura 6.40 Modo associado.

No **modo não associado** o percurso seguido pelas mensagens de sinalização entre centrais de comutação é diferente do percurso seguido pelas mensagens telefónicas, sendo o seu encaminhamento efectuado por STPs (ver Figura 6.40). Além disso, esse percurso varia ao longo do tempo consoante a disponibilidade de vias de acesso e dos STPs, podendo dizer-se que cada mensagem segue um percurso aleatório, sem nenhum trajecto previamente definido pela rede.

O modo **quase associado** é um caso limite do modo não associado. Aqui, a rede também faz recurso de STPs para fins de encaminhamento, mas o trajecto seguido pelas mensagens é pré-definido. Por exemplo, na Figura 6.41 o tráfego de voz entre as centrais A e C é transportado pela ligação directa, enquanto a correspondente sinalização é encaminhada através do nó B, que tem também funções de STP. O modo quase associado é muitas vezes usado como alternativa no caso da falha de uma ligação de sinalização. Voltando de novo à Figura 6.41 podia-se por a hipótese de as centrais A e B operarem em modo associado e fazerem recurso ao modo quase associado em caso de falha do primeiro.

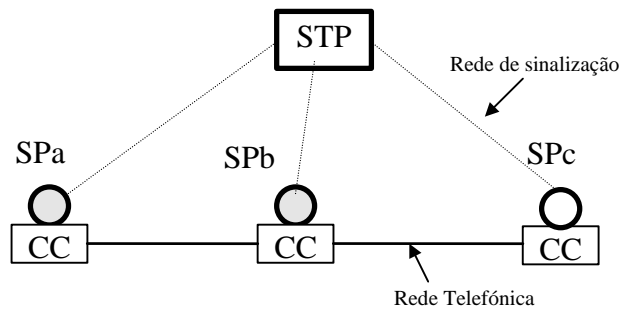



Figura 6.41 Modo não associado.

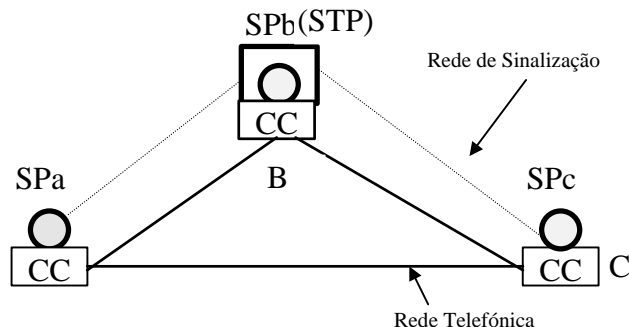



Figura 6.42 Modo quase associado.

Normalmente quando o tráfego entre duas centrais é elevado, faz-se recurso ao modo associado. Por exemplo, no caso de se ter 200 circuitos entre suas centrais, teria de se usar 7 sistemas de transmissão digital a 2 Mb/s, um dois quais seria usado para transportar a via de sinalização de canal comum no intervalo de tempo nº 16.

6.3.7.4 Vantagens e desvantagens da sinalização em canal comum

As principais **vantagens** são as seguintes:

- As duas redes (sinalização e dados) podem evoluir independentemente uma da outra.

- A informação de sinalização pode ser transmitida muito mais rapidamente entre os processadores, permitindo diminuir os tempos de estabelecimento e terminação das ligações, e deste modo aproveitar de modo mais eficiente as troncas.

- Através da ligação da rede de sinalização às bases de dado, ganha-se capacidade para introduzir toda uma nova gama de serviços, através de simples modificação do *software* armazenado nessas bases de dados.

- Os sinais de sinalização relativos a uma chamada podem ser enviados, enquanto essa chamada está a decorrer. Enviando sinais de interrogação através da rede de sinalização, a central originária pode obter informação do estado do assinante destinatário, assim como, das centrais destinatárias, antes de fazer qualquer ligação na rede telefónica. Se por exemplo, o assinante destinatário está livre, a central local estabelece um circuito de voz. Se em contrapartida está ocupado, aborta a chamada sem ter estabelecido nenhum circuito de voz.

- Podem ser transmitidos sinais de sinalização entre os processadores para outras funções distintas das relacionadas com o processamento de chamadas, por exemplo para funções de gestão e manutenção da rede.

As principais **desvantagens** podem-se sintetizar do seguinte modo:

• É necessário um grau de fiabilidade elevado na rede de sinalização, tanto em termos dos meios físicos (requerendo a duplicação dos STPs), como em termos do controlo de erros.

• Como o percurso do sinal de voz e sinalização são diferentes, não há um teste automático da integridade do circuito. É necessário introduzir sinais de teste adicionais para o circuito de voz.

6.3.7.5 Arquitectura funcional do SS7

O protocolo do Sistema #7 é estruturado em duas partes: parte do utilizador **UP** (*User Part*) e parte da transferência de mensagens **MTP** (*Message Transfer Part*). A parte de transferência de mensagens corresponde aos níveis 1 a 3 e a parte do utilizador ao nível 4. O MTP tem capacidade para suportar diversos utilizadores, tendo sido definidos três: o utilizador telefónico com lacete do assinante analógico (**TUP**), utilizador RDIS (**ISUP**) e utilizador de dados (**DUP**).

Contudo, o SS7 é usado actualmente para outras funções distintas das referidas acima, nomeadamente para consulta de bases de dados, gestão de tráfego e operações de vigilância, manutenção e gestão. Estas novas funções levaram ao desenvolvimento de uma outra parte designada por **TC** (*transaction capabilities*). Contrariamente às partes anteriores, o TC foi projectado já depois do modelo OSI ter sido especificado. É, assim, natural que, tenha sido estruturado usando esse modelo como base. Para isso, foi adicionado ao nível 3 do protocolo SS7 um outro nível o **SCCP** (*signalling -connection control part*) para o tornar totalmente compatível com o nível 3 do OSI. O **TC-ISP** (TC *intermediate part*) realiza as funções dos níveis 4 a 6 do OSI e o **TCAP** (*TC application part*) as do nível 7. Note-se que, a **OMAP** (*operations, maintenance and administration part*) também realiza funções do nível 7 do OSI.

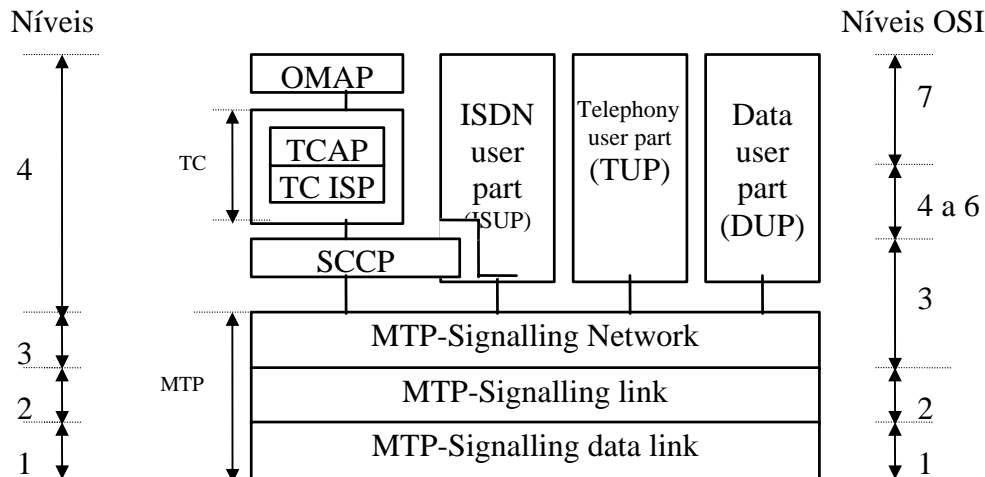



Figura 6.43 Relação entre os níveis do SS7 e do OSI.

MTP (Message Transfer Part)

O protocolo MTP tem como função a transferência fiável de mensagens de sinalização entre pares de SPs, através de vias de sinalização (linhas digitais) e STPs. Este protocolo encontra-se estruturado nos seguintes níveis:

Nível 1 (*Signalling data link*)

Define as caracteristicas físicas, eléctricas e funcionais das ligações. Por exemplo o débito binário, (o qual o ITU-T recomenda que seja igual a 64 kb/s) e o tipo de código de linha a usar na transmissão. No caso do modo de operação associado usa-se o *time-slot* 16 de uma trama de um sistema de 2 Mb/s como via de sinalização.

Nível 2 (*Signaling link*)

É responsável pela transferência de informação, estruturada em tramas que aqui se designam por SU (*signaling unit*), entre os nós da rede de sinalização adjacentes, desempenhando, nomeadamente as seguintes funções:

- Gera, transmite e recebe SUs correctamente delimitadas e sincronizadas;
- Detecta e corrige erros;
- Analisa a qualidade da transmissão, medindo a percentagem de SUs rejeitados;
- Realiza o controlo da congestão, fazendo com que o SP/STP receptor controlo o fluxo de SUs do STP emissor, etc.

As SUs podem-se dividir em **MSU** (*message signal unit*), a qual transfere a informação proveniente do nível 4, **LSSU** (*link-status signal unit*) usada para iniciar a ligação e monitorizar o seu estado, e **FISU** (*fill-in signl unit*) transmitida quando não há tráfego, no sentido do terminal receptor saber que o outro está em estado inactivo.

Uma MSU pode ser vista como uma trama de comprimento variável, que é usada para transportar as mensagens provenientes do nível 3, que são enviadas em pacotes. O formato de uma MSU está definido na Figura 6.44.

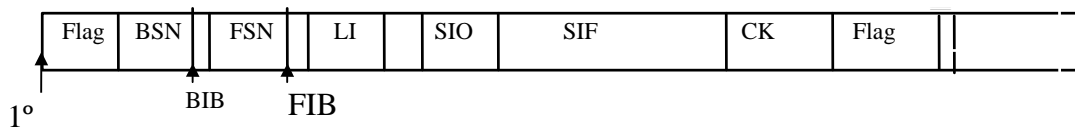



Figura 6.44 Estrutura de uma MSU.

A informacão proveniente do nível três ocupa o SIF (*signal information field*) a qual tem comprimento variável sendo constituida por *n* bytes (n≤ 272). As funcões dos outros bits são as seguintes:

- **Flags**: São usadas para delimitar as mensagens e apresentam o padrão 01111110;

- **FIB** (*forward indicator bit*), **BIB** (*backward indicator bit*), **FSN** (*forward sequence number*), e **BSN** (*backward sequence number*) são usados para a correção de erros;

O FSN é o número de identificação da MSU que está sendo transmitida, enquanto o BSN designa o número da última MSU recebida correctamente. O FIB indica que a MSU está sendo retransmitida, no caso em que o seu valor é o inverso do anterior. Se o valor do FIB da MSU anterior e da actual são o mesmo, isto quer dizer que a MSU actual está a ser transmitida pela primeira vez. Quando o BIB é inverso do anterior, foi detectado um erro na MSU recebida e requer-se a sua retransmissão. Quando se mantem o mesmo, isto quer dizer que não foi detectado qualquer erro.

• **LI** (*length indicator*): dá o comprimento da SU, permitindo deste modo espeificar o tipo de SU (LI=0, FISU, LI=1 ou 2 LISU e LI>2 MSU);

• **SIO** (*service information octet*): contém informação sobre o tipo de utilizador (Ex: telefone, dados, RDIS);

• **CK** (*check*): permite detectar erros.Quando um erro é detectado a MSU é ignorada e inicia-se um processo de correcção de erros, que implica a retransmissão da mensagem errada.

Nível 3 (*Signaling network*)

Define as funções que são necessárias numa rede com mais do que uma ligação. É, assim, responsável pelo encaminhamento das mensagens e pela verificação se as mensagens já chegaram ao destino. Neste caso, entrega-as à parte adequada do nível 4. Além disso, promove a gestão das falhas e congestionamento das diversas ligações da rede, sendo, assim, também responsável pela transferência fiável de mensagems entre centrais mesmo em caso de falha do nível 1 ou do 2.

O cabeçalho definido por este nível, de acordo com o padrão ITU-T, tem 32 bits (ver Figura 6.45).

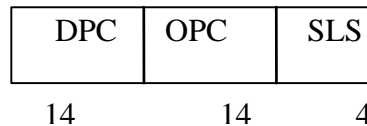



Figura 6.45 Cabeçalho introduzido pelo nível 3.

Cada ponto de sinalização dentro da rede é identificado por um código do ponto, isto é um endereço ou código de 14 bits, que é único dentro da rede. O **DPC** (*destination point code*) identifica o ponto de sinalização destinatário da mensagem e o **OPC** (*originating point code*) identifica o ponto originário da mensagem. O **SLS** (*signalling link selection*) é usado para identificar a via de sinalização usada para transferir a mensagem.

Como já se viu, por razões de fiabilidade os STS são duplicados, havendo deste modo sempre mais do que uma via de sinalização entre dois pontos de sinalização. Assim, quando uma mensagem é para ser envida o nível 3 do SP selecciona aleatoriamente um SLS, de um conjunto possível, de modo a originar uma distribuição uniforme do tráfego. Este SLS define o percurso ao longo da rede entre o SP emissário e o SP destinatário. Por exemplo, de acordo com a Figura 6.47 a seleção do SLS=xx10, conduziria ao seguinte percurso: SP A → STP A2→STP B1 →SP B.

Note-se que, para a topologia apresentada são necessários sómente dois bits do código SLS. Os outos bits poderiam ser usados ou para implementar topologias mais

complexas, ou para outras funções, como por exemplo o estabelecimento de um canal de dados adicional.

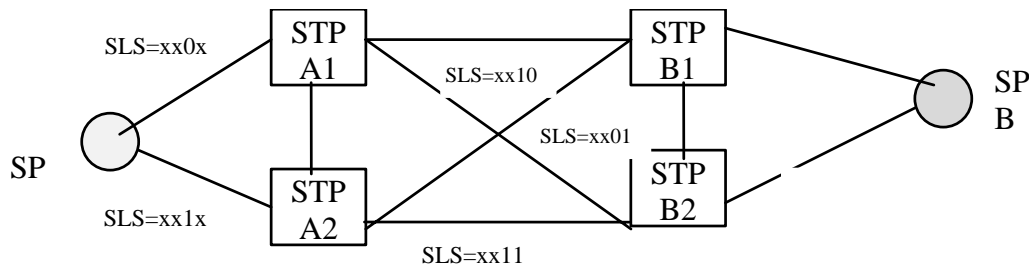



Figura 6.46 Exemplo de utilização do SLS.

Nível 4 (SCCP)

O **SCCP** foi criado para aumentar as potencialides do MTP, de modo a que juntos possam garantir ao SS7 a realização dos serviços de rede (nível 3) do modelo OSI. O objectivo do SCCP é permitir a transferência de dados entre dois nós, mesmo quando não estão envolvidos circuitos de voz, como é o caso do acesso a base de dados. Proporciona, nomedamente, controlo do fluxo e controlo da sequência das mensagens. Possibilita ainda a transferência de mensagens, quer usando datagramas, quer circuitos virtuais.

Nível 4 (ISUP)

O **ISUP** (*ISDN User Part*) é um protocolo que proporciona as funções básicas de controlo de chamadas ISDN para o estabelecimento, manutenção e terminação de ligações comutadas na rede telefónica. Pode interagir tanto com o SCCP como com o MTP, dependendo do nível do serviço exigido para o transporte das suas mensagens de sinalização.

Nivel 4 (TUP)

O **TUP** (*Telephone user part*), que é o antecessor do ISUP suporta unicamente ligações de voz usando linhas analógicas.

Nível 4 (DUP)

O **DUP** (*data user part*) define o protocolo para realizar serviços comutados de dados.

Nível 4 (TC)

Suporte de aplicações especiais não ligadas a chamadas telefónicas, como por exemplo o acesso a bases de dados, gestão do trafego, etc. É constituído por duas partes o **ISP** (*Intermediate Service Part*) que realiza as funções 4 a 6 do modelo OSI e o **TCAP** (*Transaction Capability Application Part*) que proporciona as funções 7.

O ISP é requerido quando o TC é baseado no serviço que usa circuitos virtuais. Quando o serviço usa datagramas as funções do ISP não são necessárias, e os níveis de 4 a 6 podem ser considerados como transparentes. O TCAP é constituído pela sub-camada de transacção (*transation sub-layer*) e na sub-camada de componente (*component sub-layer*).

A primeita é responsável por estabelecer e manter a ligação entre dois pontos de sinalização, enquanto a segunda é responsável por iniciar uma acção num nó remoto e retornar os resultados dessa acção.

Nível 4 (OMAP)

Responsável por funções de operação, manutenção e administração da rede.

Na Figura 5.47 apresenta-se um exemplo de aplicação dos diferentos elementos da arquitectura do SS7. Nessa figura duas centrais de comutação telefónica estão ligadas por vários circuitos de voz (a cheio). Considere-se, como exemplo que o ISUP da central A recebe um pedido para estabecer um circuito de voz . A central A reconhece que essa chamada requer informação de encaminhamento especial. Então o sistema de controlo dessa central pede ao TC para obter a partir de uma base de dados de um SPC a informação necessária. Essa função é desempenhada usando o SCCP e o MTP sem o estabelecimento de nenhum circuito de voz para esse SCP.

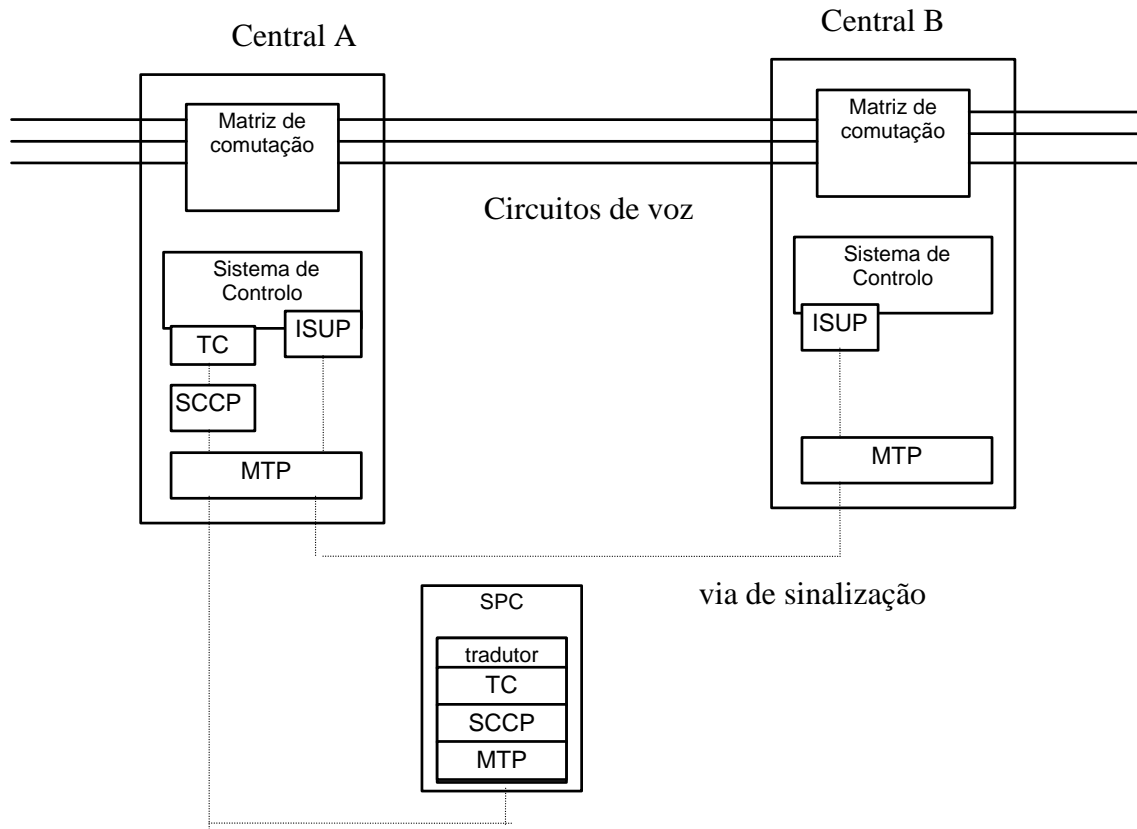



Figura 6.47 Exemplo de aplicação de elementos da arquitectura do SS7.

O tradutor da base de dados fornece a informação necessária e envia-a para o sistema de controlo da central A, através do MTP, SCCP e TC. Usando esta informação o sistema de controlo da Central A pede ao ISUP para estabelecer um circuito de voz da central A para

a central B. Isto é conseguido através da comunicação apropriada entre os ISUP das duas centrais.

6.3.8 Referências

[Fe87] K. Feher, *Advanced Digital Communications*, Prentice-Hall, 1987.

[Fl95] J. E. Flood*, Telecommunications Switching, Trafic and Networks*, Prentice- Hall*,* 1995.

[La95] F. Lavrador "Sobre as redes de telecomunicações ditas "Inteligentes" e a evolução a que a elas conduziu" *Ingenium*, pág. 96-102, Setembro de 1995.

[Ma91] R. Manterfield, *Commom-Channel Signalling*, Peter Peregrinus, 1991.

[Ra94] T. Ramteke, *Networks,* Prentice-Hall, 1994.